PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/19, C07K 14/52, C12N 15/12, C07K 14/715, 16/24, 16/28, C12N 15/62, G01N 33/68, A61K 38/19

A2

(11) International Publication Number: **WO 98/32858**

(43) International Publication Date: 30 July 1998 (30.07.98)

(21) International Application Number: PCT/US98/00902

(22) International Filing Date: 22 January 1998 (22.01.98)

(30) Priority Data:
60/036,715 23 January 1997 (23.01.97) US

(71) Applicant: SCHERING CORPORATION [US/US]; 2000 Galloping Hill Road, Kenilworth, NJ 07033–0530 (US).

(72) Inventors: MATTSON, Jeanine, D.; 559 Alvarado Street, San Francisco, CA 94114 (US). SOTO–TREJO, Hortensia; Apartment 10, 1255 Cortez Drive, Sunnyvale, CA 94086 (US). HEDRICK, Joseph, A.; 52–08 Quail Ridge Drive, Plainsboro, NJ 08536 (US). GORMAN, Daniel, M.; 6371 Central Avenue, Newark, CA 94560 (US). ZLOTNIK, Albert; 507 Alger Drive, Palo Aito, CA 94306 (US).

(74) Agents: McLAUGHLIN, Jaye, P. et al.; Schering–Plough Corporation, Patent Dept., K–6–1, 1990, 2000 Galloping Hill Road, Kenilworth, NJ 07033–0530 (US).

(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CZ, EE, GE, GW, HU, ID, IL, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MAMMALIAN CHEMOKINES; RECEPTORS; REAGENTS; USES

(57) Abstract

Novel chemokines and 7 transmembrane receptors from mammals, reagents related thereto, including purified proteins, specific antibodies, and nucleic acids encoding said chemokines or receptors. Methods of using said reagents and diagnostic kits are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MAMMALIAN CHEMOKINES; RECEPTORS; REAGENTS; USES

The present filing claims priority to U.S. Patent Application No. 60/036,715, filed January 23, 1997, which is incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to compositions related to proteins which function in controlling physiology, development, and/or differentiation of mammalian cells. In particular, it provides proteins which are implicated in the regulation of physiology, development, differentiation, or function of various cell types, e.g., chemokines, 7 transmembrane receptors, reagents related to each, e.g., antibodies or nucleic acids encoding them, and uses thereof.

BACKGROUND OF THE INVENTION

The circulating component of the mammalian circulatory system comprises various cell types, including red and white blood cells of the erythroid and myeloid cell lineages. See, e.g., Rapaport (1987) Introduction to Hematology (2d ed.) Lippincott, Philadelphia, PA; Jandl (1987) Blood: Textbook of Hematology, Little, Brown and Co., Boston, MA.; and Paul (ed.) (1993) Fundamental Immunology (3d ed.) Raven Press, N.Y.

For some time, it has been known that the mammalian immune response is based on a series of complex cellular interactions, called the "immune network." Recent research has provided new insights into the inner workings of this network. While it remains clear that much of the response does, in fact, revolve around the network-like interactions of lymphocytes, macrophages, granulocytes, and other cells, immunologists now generally hold the opinion that soluble proteins, known as lymphokines, cytokines, or monokines, play a critical role in controlling these cellular interactions. Thus, there is considerable interest in the isolation, characterization, and mechanisms of action

of cell modulatory factors, an understanding of which should lead to significant advancements in the diagnosis and therapy of numerous medical abnormalities, e.g., immune system and other disorders.

Lymphokines apparently mediate cellular activities in a variety of ways. They have been shown to support the proliferation, growth, and differentiation of the pluripotential hematopoietic stem cells into vast numbers of progenitors comprising diverse cellular lineages making up a complex immune system. These interactions between the cellular components are necessary for a healthy immune response. These different cellular lineages often respond in a different manner when lymphokines are administered in conjunction with other agents.

The chemokines are a large and diverse superfamily of proteins. The superfamily is subdivided into two classical branches, based upon whether the first two cysteines in the chemokine motif are adjacent (termed the "C-C" branch), or spaced by an intervening residue ("C-X-C"). A more recently identified branch of chemokines lacks two cysteines in the corresponding motif, and is represented by the chemokines known as lymphotactins. Another recently identified branch has three intervening residues between the two cysteines, e.g., CX3C chemokines. See, e.g., Schall and Bacon (1994) Current Opinion in Immunology 6:865-873; and Bacon and Schall (1996) Int. Arch. Allergy & Immunol. 109:97-109.

The chemokine receptors are typically members of the superfamily of G-protein coupled (or linked) receptors (GPCR, or GPLR). As a class, these receptors are integral membrane proteins characterized by amino acid sequences which contain seven hydrophobic domains. See, e.g., Ruffolo and Hollinger (eds. 1995) G-Protein Coupled Transmembrane Signaling Mechanisms CRC Press, Boca Raton, FL; Watson and Arkinstall (1994) The G-Protein Linked Receptor FactsBook Academic Press, San Diego, CA; Peroutka (ed. 1994) G Protein-Coupled Receptors CRC Press, Boca Raton, FL; Houslay and Milligan (1990) G-Proteins as Mediators of Cellular Signaling Processes Wiley and Sons, New York, NY; and Dohlman, et al. (1991) Ann. Rev. Biochem. 60:653-688. These hydrophobic

domains are predicted to represent transmembrane spanning regions of the proteins. These GPCRs are found in a wide range of organisms and are typically involved in the transmission of signals to the interior of the cell, e.g., through interaction, e.g., with heterotrimeric G-proteins. They respond to a wide and diverse range of agents including lipid analogs, amino acid derivatives, small peptides, and other molecules.

The presumed transmembrane segments are typically 20-25 amino acids in length. Based upon models and data on bacteriorhodopsin, these regions are predicted to be a-helices and be oriented to form a ligand binding pocket. See, e.g., Findley, et al. (1990) Trends Pharmacol. Sci. 11:492-499. Other data suggest that the amino termini of the proteins are extracellular, and the carboxy termini are intracellular. See, e.g., Lodish, et al. (1995) Molecular Cell Biology 3d ed., Scientific American, New York; and Watson and Arkinstall (1994) The G-Protein Linked Receptor FactsBook Academic Press, San Diego, CA. Phosphorylation cascades have been implicated in the signal transduction pathway of these receptors.

Although the full spectrum of biological activities mediated by these 7 transmembrane receptors has not been fully determined, chemoattractant effects are recognized. Chemokine receptors are notable members of the GPCR family. See, e.g., Samson, et al. (1996) Biochemistry 35:3362-3367; and Rapport, et al. (1996) J. Leukocyte Biology 59:18-23. The best known biological functions of these chemokine molecules relate to chemoattraction of leukocytes. However, new chemokines and receptors are being discovered, and their biological effects on the various cells responsible for immunological responses are topics of continued study.

Many factors have been identified which influence the differentiation process of precursor cells, or regulate the physiology or migration properties of specific cell types. These observations indicate that other factors exist whose functions in immune function were heretofore unrecognized. These factors provide for biological activities whose spectra of effects may be distinct from known differentiation or activation factors. The absence of knowledge about the structural, biological, and physiological

properties of the regulatory factors which regulate cell physiology in vivo prevents the modulation of the effects of such factors.

In addition, other factors exist whose functions in hematopoiesis, neural function, immune development, and leukocyte trafficking were heretofore unrecognized. These receptors mediate biological activities whose spectra of effects are distinct from known differentiation, activation, or other signaling factors. The absence of knowledge about the structural, biological, and physiological properties of the receptors which regulate cell physiology, development, or function prevents the modification of the effects of such factors.

Thus, medical conditions where regulation of the development or physiology of relevant cells is required remain unmanageable.

SUMMARY OF THE INVENTION

The present invention is based, in part, upon the discovery of new genes encoding various chemokines, e.g., those designated IBICK, which encodes primate CXC chemokines; ILINCK, which encodes primate CXC chemokines; CXC-143, which encodes rodent CXC chemokines; MCP243, which encodes a mouse chemokine; or 7 transmembrane receptors, e.g., those designated R277, which encode primate receptors; HST01.1, which encode rodent receptors; and 941D12, which encode rodent receptors. Each GPCR gene encodes a polypeptide exhibiting structural and/or sequence homology to 7 transmembrane receptors. Such receptors are typically G-protein coupled (or linked) receptors (GPCR or GPLR), though a ligand for each has not yet been identified.

The invention also provides mutations (muteins) of the respective natural sequences, fusion proteins, chemical mimetics, antibodies, and other structural or functional analogs. It is also directed to isolated nucleic acids, e.g., genes encoding respective proteins of the invention. Various uses of these different protein, antibody, or nucleic acid compositions are also provided.

The present invention provides a composition selected from the group of: a substantially pure antigenic polypeptide comprising sequence from an IBICK; an ILINCK; a CXC-143; an MCP243; an R277; an HST01.1; or a 941D12; a binding composition comprising an antigen binding portion of an antibody specific for binding to such an antigenic polypeptide; a nucleic acid encoding such an antigenic polypeptide; and a fusion protein comprising at least two non-overlapping segments of at least 10 amino acids of such an antigenic polypeptide.

In certain embodiments of the antigenic polypeptide, it is from a warm blooded animal, e.g., a rodent or primate; it comprises a sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26; it exhibits a post-translational modification pattern distinct from a natural form of said polypeptide; it is detectably labeled; or it is made by expression of a recombinant nucleic acid. In other embodiments, a sterile form is provided, including, e.g., composition comprising the polypeptide and an acceptable carrier. A detection kit comprising

a compartment or container holding such an antigenic polypeptide is also provided.

In other binding composition forms, e.g., antibody embodiments, the polypeptide is a mouse or human protein; the antibody is raised against a peptide sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26; the antibody is a monoclonal antibody; the binding composition is fused to a heterologous protein, or is detectably labeled. An alternative embodiment is a binding compound comprising an antigen binding fragment of the antibody described. Also provided is a detection kit comprising such a binding compound. With the antibodies are provided methods of purifying a polypeptide using the binding compound or antibody to specifically separate the polypeptides from others, or for detection, e.g., immunohistochemistry or immunoprecipitation.

Nucleic acid embodiments are provided, e.g., where the nucleic acid is in an expression vector and: encodes a polypeptide from a mouse or human; encodes a mature protein of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26; or comprises a deoxyribonucleic acid nucleotide. The invention also provides a kit with such nucleic acids, e.g., which include PCR primers for amplifying such sequences.

With nucleic acids are provided fusion proteins, comprising: a sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26; and/or sequence of another chemokine or 7 transmembrane receptor, as appropriate. Also provided is a cell comprising a recombinant nucleic acid, as described, and methods of producing a polypeptide comprising expressing the nucleic acid in an expression system.

Other embodiments include methods of modulating physiology or development of a cell, with a step of contacting that cell with a composition comprising an agonist or antagonist of the chemokine or receptor. Ordinarily, the cell is a neuron, macrophage, or lymphocyte. Various physiological effects to be modulated include a cellular calcium flux, a chemoattractant response, cellular morphology modification responses, phosphoinositide lipid turnover, or an antiviral response.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. General

The present invention provides DNA sequences encoding various mammalian proteins, including chemokines, or which exhibit structural properties characteristic of a 7 transmembrane receptor. See, e.g., Ruffolo and Hollinger (eds. 1995) G-Protein Coupled Transmembrane Signaling Mechanisms CRC Press, Boca Raton, FL; Watson and Arkinstall (1994) The G-Protein Linked Receptor FactsBook Academic Press, San Diego, CA; Peroutka (ed. 1994) G Protein-Coupled Receptors CRC Press, Boca Raton, FL; Houslay and Milligan (1990) G-Proteins as Mediators of Cellular Signaling Processes Wiley and Sons, New York, NY. Certain human and mouse embodiments are described herein.

Among the many types of ligands which mediate biology via these receptors are chemokines and certain proteases. Chemokines play an important role in immune and inflammatory responses by inducing migration and adhesion of leukocytes. See, e.g., Schall (1991) Cytokine 3:165-183; and Thomson (ed.) The Cytokine Handbook Academic Press, NY. Chemokines are secreted by activated leukocytes and act as a chemoattractant for a variety of cells which are involved in inflammation. Besides chemoattractant properties, chemokines have been shown to induce other biological responses, e.g., modulation of second messenger levels such as Ca^{++}; inositol phosphate pool changes (see, e.g., Berridge (1993) Nature 361:315-325 or Billah and Anthes (1990) Biochem. J. 269:281-291); cellular morphology modification responses; phosphoinositide lipid turnover; possible antiviral responses; and others. Thus, the chemokines provided herein may, alone or in combination with other therapeutic reagents, have advantageous combination effects.

Moreover, there are reasons to suggest that chemokines may have effects on other cell types, e.g., attraction or activation of monocytes, dendritic cells, T cells, eosinophils, and/or perhaps on basophils and/or neutrophils. They may also have chemoattractive effects on various neural cells including, e.g., dorsal root ganglia

neurons in the peripheral nervous system and/or central nervous system neurons.

G-protein coupled receptors, e.g., chemokine receptors, are important in the signal transduction mechanisms mediated by their ligands. They are useful markers for distinguishing cell populations, and have been implicated as specific receptors for retroviral infections.

The chemokine superfamily was classically divided into two groups exhibiting characteristic structural motifs, the Cys-X-Cys (C-X-C) and Cys-Cys (C-C) families. These were distinguished on the basis of a single amino acid insertion between the NH-proximal pair of cysteine residues and sequence similarity. Typically, the C-X-C chemokines, i.e., IL-8 and MGSA/Gro-a act on neutrophils but not on monocytes, whereas the C-C chemokines, i.e., MIP-1a and RANTES, are potent chemoattractants for monocytes and lymphocytes but not neutrophils. See, e.g., Miller, et al. (1992) Crit. Rev. Immunol. 12:17-46. A recently isolated chemokine, lymphotactin, does not belong to either group and may constitute a first member of a third chemokine family, the C family. Lymphotactin does not have a characteristic CC or CXC motif, and acts on lymphocytes but not neutrophils and monocytes. See, e.g., Kelner et al. (1994) Science 266:1395-1399. This chemokine defines a new C-C chemokine family. Even more recently, another chemokine exhibiting a CX3C motif has been identified, which establishes a fourth structural class.

The present invention provides additional chemokine reagents, e.g., nucleic acids, proteins and peptides, antibodies, etc., related to the newly discovered chemokines designated primate IBICK; primate ILINCK; rodent CXC-143; or rodent MCP243.

In other embodiments, the invention provides genes encoding novel G-protein coupled receptors, designated primate R277, rodent HST01.1, and rodent 941D12. Their ligands have not yet specifically been identified. However, the receptors exhibit structural features typical of known 7 transmembrane spanning receptors, which receptors include chemokine receptors. The receptors may exhibit properties of binding many different cytokines

at varying specificities (shared or promiscuous binding specificity) or may exhibit high affinity for one (specific) or a subset (shared) of chemokines. Alternatively, the ligands may be other molecules, including molecules such as epinephrine, serotonin, or glucagon.

The described chemokines or receptors should be important for mediating various aspects of cellular, organ, tissue, or organismal physiology or development.

II. Purified Chemokines; Receptors

Nucleotide and derived amino acid sequences of a human embodiment of a primate CXC chemokine, designated IBICK are shown in SEQ ID NO: 1 and 2. The term "IBICK" will encompass other primate counterparts. The gene encodes a novel protein exhibiting structure and motifs characteristic of a chemokine. The predicted signal cleavage site is around the gly(-1)-phe1 peptide bond. Complementary nucleic acid sequences may be used for many purposes, e.g., in a PCR primer pair or as a mutagenesis primer. Fragments of the nucleotide sequence may be used as hybridization probes, or PCR primers, or to encode antigenic peptides. Fragments of the polypeptide will be useful as antigenic peptides. Likewise for the other genes. The closest reported chemokines to the IBICK sequences are the MIG and IP10 chemokines, both of which are IFN-γ regulated. See, e.g., Faubert (1993) Biochem. Biophys. Res. Commun. 192:223-230; and Luster, et al. (1985) Nature 315:672-676.

Nucleotide and derived amino acid sequences of a novel primate CC chemokine, e.g., from human, designated ILINCK are shown in SEQ ID NO: 3 and 4. The term "ILINCK" as used in this filing will encompass other primate counterparts. The predicted signal cleavage site is around the ser(-1)-gln1 peptide bond. Two different messages have been detected which encode the chemokine, and the larger one, a 1.5 kB message, is upregulated by IL-10. This is an unusual property of chemokine messages, which implies that the chemokine has a role in anti-inflammatory responses.

Partial nucleotide and derived amino acid sequences of a novel rodent CXC chemokine, e.g., from mouse, designated CXC-143, are shown in SEQ ID NO: 5, 6, 7, 8, 9 and 10. The term "CXC-143"

will typically encompass rodent counterparts. Sequence analysis shows closest sequence homology to the IP10, MIG, and IBICK chemokines, described above. It may well be that the chemokine will be similarly regulated, e.g., by IFN-γ.

Nucleotide and derived amino acid sequences of a novel rodent chemokine, e.g., from mouse, designated MCP243 are shown in SEQ ID NO: 11, 12, 13 and 14. The term may encompass other rodent counterparts. Cys residues 14 and 30 correspond to conserved chemokine Cys3 and Cys4. Sequence analysis shows closest sequence homology to other chemokines.

Nucleotide and derived amino acid sequences of a novel rodent GPCR, e.g., from mouse, designated R277, are shown in SEQ ID NO: 15, 16, 17 and 18. The term "R277" may encompass other primate counterparts. Note that nucleotide 447 is designated C, but may be C or T; nucleotides 489 and 640 are each designated C, but may be A, C, G, or T; and nucleotides 480-510 may contain various sequence errors, but which will retain reading frame.. Sequence analysis shows closest sequence homology to a human GPCR designated TDAG8.

Partial nucleotide and derived amino acid sequences of a novel rodent GPCR, e.g., from mouse, designated HST01.1, are shown in SEQ ID NO: 19 and 20. The term "R277" may encompass other primate counterparts. Note that nucleotide 447 is designated C, but may be C or T; nucleotides 489 and 640 are each designated C, but may be A, C, G, or T; and nucleotides 480-510 may contain various sequence errors, but which will retain reading frame. The sequence is supplemented with more complete sequence in SEQ ID NO: 21 and 22. A DRY box motif runs from about asp147 to ala155; transmembrane segments run from about ala57 to leu78; phe90 to val110; val125 to phe146; val167 to leu189; phe223 to val243; leu255 to leu279; and val301 to val322. Sequence analysis shows sequence homology to various GPCR family members.

Partial nucleotide and derived amino acid sequences of a novel rodent GPCR, e.g., from mouse, designated 941D12, are shown in SEQ ID NO: 23 and 24. The term "941D12" may encompass other rodent counterparts. The nucleotides at positions 169, 178, 217, 287,

290, 382, 386, 395, 411, 484, 512, 515, 517, and 521 are each indicated as C, but may be A, C, G, or T. A complete rodent 941D12 is provided in SEQ ID NO: 25 and 26. Nucleotide 942 is designated C, but may be C or T; nucleotides 1412 and 1422 each designated C, but may be A, C, G, or T. Sequence analysis shows sequence homology to various GPCR family members.

Certain general descriptions of physical properties of polypeptides, nucleic acids, and antibodies, where directed to one embodiment clearly are usually applicable to other chemokines or receptors described herein.

These amino acid sequences, provided amino to carboxy, are important in providing sequence information on the chemokine ligand or receptor, allowing for distinguishing the protein from other proteins, particularly naturally occurring versions. Moreover, the sequences allow preparation of peptides to generate antibodies to recognize and distinguish such segments, and allow preparation of oligonucleotide probes, both of which are strategies for isolation, e.g., cloning, of genes encoding such sequences, or related sequences, e.g., natural polymorphic or other variants, including fusion proteins. Similarities of the chemokines have been observed with other cytokines. See, e.g., Bosenberg, et al. (1992) Cell 71:1157-1165; Huang, et. al. (1992) Molecular Biology of the Cell 3:349-362; and Pandiella, et al. (1992) J. Biol. Chem. 267:24028-24033. Likewise for the GPC receptors.

As used herein, the term "IBICK" shall encompass, when used in a protein context, a protein having mature amino acid sequence, as shown in SEQ ID NO: 2. The invention also embraces a polypeptide comprising a significant fragment of such protein. The invention also encompasses a polypeptide which is a primate species counterpart, e.g., which exhibits similar sequence, and is more homologous in natural encoding sequence than other genes from a primate species. Typically, such chemokine will also interact with its specific binding components, e.g., receptor, or antibodies which bind to it. These binding components, e.g., antibodies, typically bind to the chemokine with high affinity, e.g., at least about 100 nM, usually better than about 30 nM, preferably better than about 10 nM, and

more preferably at better than about 3 nM. Similar concepts apply to the primate embodiments for the chemokine ILINCK and the GPCR R277. In contrast, rodent embodiments for the chemokines CXC-143 and MCP243, and the GPCRs HST01.1 and 941D12 encompass other rodent species counterparts.

The term "polypeptide" as used herein includes a significant fragment or segment, and encompasses a stretch of amino acid residues of at least about 8 amino acids, generally at least 10 amino acids, more generally at least 12 amino acids, often at least 14 amino acids, more often at least 16 amino acids, typically at least 18 amino acids, more typically at least 20 amino acids, usually at least 22 amino acids, more usually at least 24 amino acids, preferably at least 26 amino acids, more preferably at least 28 amino acids, and, in particularly preferred embodiments, at least about 30 or more amino acids, e.g., about 35, 40, 45, 50, 60, 75, 80, 100, 120, etc. Similar proteins will likely comprise a plurality of such segments. Such fragments may have ends which begin and/or end at virtually all positions, e.g., beginning at residues 1, 2, 3, etc., and ending at, e.g., 69, 68, 67, 66, etc., in all combinatorial pairs in the coding segment. Particularly interesting peptides have ends corresponding to structural domain boundaries, e.g., intracellular or extracellular loops of the receptor embodiments. Such peptides will typically be immunogenic peptides, or may be concatenated to generate larger polypeptides. Short peptides may be attached or coupled to a larger carrier.

The term "binding composition" refers to molecules that bind with specificity to the respective chemokine or receptor, e.g., in a ligand-receptor type fashion or an antibody-antigen interaction. These compositions may be compounds, e.g., proteins, which specifically associate with the chemokine or receptor, including natural physiologically relevant protein-protein interactions, either covalent or non-covalent. The binding composition may be a polymer, or another chemical reagent. No implication as to whether the chemokine presents a concave or convex shape in its ligand-receptor interaction is necessarily represented, other than the interaction exhibit similar specificity, e.g., specific affinity. A functional analog may be a ligand with structural modifications, or

may be a wholly unrelated molecule, e.g., which has a molecular shape which interacts with the appropriate ligand binding determinants. The ligands may serve as agonists or antagonists of a physiological or natural receptor, see, e.g., Goodman, et al. (eds.) (1990) Goodman & Gilman's: The Pharmacological Bases of Therapeutics (8th ed.), Pergamon Press. The term expressly includes antibodies, polyclonal or monoclonal, which specifically bind to the respective antigen.

Substantially pure means that the protein is free from other contaminating proteins, nucleic acids, and/or other biologicals typically derived from the original source organism. Purity may be assayed by standard methods, and will ordinarily be at least about 40% pure, more ordinarily at least about 50% pure, generally at least about 60% pure, more generally at least about 70% pure, often at least about 75% pure, more often at least about 80% pure, typically at least about 85% pure, more typically at least about 90% pure, preferably at least about 95% pure, more preferably at least about 98% pure, and in most preferred embodiments, at least 99% pure. Analyses will typically be by weight, but may be by molar amounts.

Solubility of a polypeptide or fragment depends upon the environment and the polypeptide. Many parameters affect polypeptide solubility, including temperature, electrolyte environment, size and molecular characteristics of the polypeptide, and nature of the solvent. Typically, the temperature at which the polypeptide is used ranges from about 4° C to about 65° C. Usually the temperature at use is greater than about 18° C and more usually greater than about 22° C. For diagnostic purposes, the temperature will usually be about room temperature or warmer, but less than the denaturation temperature of components in the assay. For therapeutic purposes, the temperature will usually be body temperature, typically about 37° C for humans, though under certain situations the temperature may be raised or lowered in situ or in vitro.

The electrolytes will usually approximate in situ physiological conditions, but may be modified to higher or lower ionic strength where advantageous. The actual ions may be modified, e.g., to

conform to standard buffers used in physiological or analytical contexts.

The size and structure of the polypeptide should generally be in a substantially stable state, and usually not in a denatured state, though in certain circumstances denatured protein will be important. The polypeptide may be associated with other polypeptides in a quaternary structure, e.g., to confer solubility, or associated with lipids or detergents in a manner which approximates natural lipid bilayer interactions.

The solvent will usually be a biologically compatible buffer, of a type used for preservation of biological activities, and will usually approximate a physiological solvent. Usually the solvent will have a neutral pH, typically at least about 5, preferably at least 6, and typically less than 10, preferably less than 9, and more preferably about 7.5. On some occasions, a detergent will be added, typically a mild non-denaturing one, e.g., CHS (cholesteryl hemisuccinate) or CHAPS (3-([3-cholamido-propyl]dimethylammonio)-1-propane sulfonate), or a low enough concentration as to avoid significant disruption of structural or physiological properties of the protein.

Solubility is reflected by sedimentation measured in Svedberg units, which are a measure of the sedimentation velocity of a molecule under particular conditions. The determination of the sedimentation velocity was classically performed in an analytical ultracentrifuge, but is typically now performed in a standard ultracentrifuge. See, Freifelder (1982) Physical Biochemistry (2d ed.), W.H. Freeman; and Cantor and Schimmel (1980) Biophysical Chemistry, parts 1-3, W.H. Freeman & Co., San Francisco. As a crude determination, a sample containing a putatively soluble polypeptide is spun in a standard full sized ultracentrifuge at about 50K rpm for about 10 minutes, and soluble molecules will remain in the supernatant. A soluble particle or polypeptide will typically be less than about 30S, more typically less than about 15S, usually less than about 10S, more usually less than about 6S, and, in particular embodiments, preferably less than about 4S, and more preferably less than about 3S.

III. Physical Variants

This invention also encompasses proteins or peptides having substantial amino acid sequence homology with the amino acid sequence of each respective receptor. The variants include species or polymorphic variants.

Amino acid sequence homology, or sequence identity, is determined by optimizing residue matches, if necessary, by introducing gaps as required. This changes when considering conservative substitutions as matches. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Homologous amino acid sequences are typically intended to include natural allelic and interspecies variations in each respective protein sequence. Typical homologous proteins or peptides will have from 25-100% homology (if gaps can be introduced), to 50-100% homology (if conservative substitutions are included) with the amino acid sequence of the appropriate chemokine or receptor. Homology measures will be at least about 35%, generally at least 40%, more generally at least 45%, often at least 50%, more often at least 55%, typically at least 60%, more typically at least 65%, usually at least 70%, more usually at least 75%, preferably at least 80%, and more preferably at least 80%, and in particularly preferred embodiments, at least 85% or more. See also Needleham, et al. (1970) J. Mol. Biol. 48:443-453; Sankoff, et al. (1983) Chapter One in Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison Addison-Wesley, Reading, MA; and software packages from IntelliGenetics, Mountain View, CA; and the University of Wisconsin Genetics Computer Group, Madison, WI.

Each of the isolated chemokine or GPC receptor DNAs can be readily modified by nucleotide substitutions, nucleotide deletions, nucleotide insertions, and inversions of nucleotide stretches. These modifications may result in novel DNA sequences which encode these antigens, their derivatives, or proteins having similar physiological, immunogenic, or antigenic activity. These modified

sequences can be used to produce mutant antigens or to enhance expression, or to introduce convenient enzyme recognition sites into the nucleotide sequence without significantly affecting the encoded protein sequence. Enhanced expression may involve gene amplification, increased transcription, increased translation, and other mechanisms. Such mutant receptor derivatives include predetermined or site-specific mutations of the respective protein or its fragments. "Mutant chemokine" encompasses a polypeptide otherwise falling within the homology definition of the chemokine as set forth above, but having an amino acid sequence which differs from that of the chemokine as found in nature, whether by way of deletion, substitution, or insertion. Likewise for the GPCRs. These include amino acid residue substitution levels from none, one, two, three, five, seven, ten, twelve, fifteen, etc. In particular, "site specific mutant" generally includes proteins having significant homology with a protein having sequences of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26, and as sharing various biological activities, e.g., antigenic or immunogenic, with those sequences, and in preferred embodiments contain most of the disclosed sequences, particularly those found in various groups of animals. As stated before, it is emphasized that descriptions are generally meant to encompass the various chemokine or receptor proteins from other members of related groups, not limited to the mouse or human embodiments specifically discussed.

Although site specific mutation sites are often predetermined, mutants need not be site specific. Chemokine or receptor mutagenesis can be conducted by making amino acid insertions or deletions. Substitutions, deletions, insertions, or combinations may be generated to arrive at a final construct. Insertions include amino- or carboxy- terminal fusions. Random mutagenesis can be conducted at a target codon and the expressed mutants can then be screened for the desired activity. Methods for making substitution mutations at predetermined sites in DNA having a known sequence are well known in the art, e.g., by M13 primer mutagenesis or polymerase chain reaction (PCR) techniques. See also Sambrook, et al. (1989) and Ausubel, et al. (1987 and Supplements). Many

structural features are known about the chemokines and GPCRs which allow determination of whether specific residues are embedded into the core of the secondary or tertiary structures, or whether the residues will have relatively little effect on protein folding. Preferred positions for mutagenesis are those which do not prevent functional folding of the resulting protein.

The mutations in the DNA normally should not place coding sequences out of reading frames and preferably will not create complementary regions that could hybridize to produce secondary mRNA structure such as loops or hairpins. But certain situations exist where such problems are compensated. See, e.g., Gesteland and Atkins (1996) Ann. Rev. Biochem. 65:741-768.

The present invention also provides recombinant proteins, e.g., heterologous fusion proteins using segments from these proteins, or antibodies. A heterologous fusion protein is a fusion of proteins or segments which are naturally not normally fused in the same manner. Thus, the fusion product of an immunoglobulin with a receptor polypeptide is a continuous protein molecule having sequences fused in a typical peptide linkage, typically made as a single translation product and exhibiting properties derived from each source peptide. A similar chimeric concept applies to heterologous nucleic acid sequences.

In addition, new constructs may be made from combining similar functional or structural domains from other proteins. For example, ligand-binding or other segments may be "swapped" between different new fusion polypeptides or fragments. See, e.g., Cunningham, et al. (1989) Science 243:1330-1336; and O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992. Thus, new chimeric polypeptides exhibiting new combinations of specificities will result from the functional linkage of ligand-binding specificities and other functional domains. Such may be chimeric molecules with mixing or matching of the various structural segments, e.g., the β-sheet or α-helix structural domains for the chemokine, or receptor segments corresponding to each of the transmembrane segments (TM1-TM7), or the intracellular (cytosolic, C1-C4) or extracellular (E1-E4) loops

from the various receptor types. The C3 loop is particularly important.

The phosphoramidite method described by Beaucage and Carruthers (1981) Tetra. Letts. 22:1859-1862, will produce suitable synthetic DNA fragments. A double stranded fragment will often be obtained either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence, e.g., PCR techniques.

IV. Functional Variants

The blocking of physiological response to various embodiments of these chemokines or GPCRs may result from the inhibition of binding of the ligand to its receptor, likely through competitive inhibition. Thus, in vitro assays of the present invention will often use isolated protein, membranes from cells expressing a recombinant membrane associated receptor, e.g., ligand binding segments, or fragments attached to solid phase substrates. These assays will also allow for the diagnostic determination of the effects of either binding segment mutations and modifications, or ligand mutations and modifications, e.g., ligand analogs.

This invention also contemplates the use of competitive drug screening assays, e.g., where neutralizing binding compositions, e.g., antibodies, to antigen or receptor fragments compete with a test compound for binding to the protein. In this manner, the antibodies can be used to detect the presence of polypeptides which share one or more antigenic binding sites of the ligand and can also be used to occupy binding sites on the protein that might otherwise interact with a receptor.

Additionally, neutralizing antibodies against a specific chemokine embodiment and soluble fragments of the chemokine which contain a high affinity receptor binding site, can be used to inhibit chemokine activity in tissues, e.g., tissues experiencing abnormal physiology.

"Derivatives" of chemokine antigens include amino acid sequence mutants, glycosylation variants, and covalent or aggregate

conjugates with other chemical moieties. Covalent derivatives can be prepared by linkage of functionalities to groups which are found in chemokine amino acid side chains or at the N- or C- termini, by means which are well known in the art. These derivatives can include, without limitation, aliphatic esters or amides of the carboxyl terminus, or of residues containing carboxyl side chains, O-acyl derivatives of hydroxyl group-containing residues, and N-acyl derivatives of the amino terminal amino acid or amino-group containing residues, e.g., lysine or arginine. Acyl groups are selected from the group of alkyl-moieties including C3 to C18 normal alkyl, thereby forming alkanoyl aroyl species. Covalent attachment to carrier proteins may be important when immunogenic moieties are haptens.

In particular, glycosylation alterations are included, e.g., made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing, or in further processing steps. Particularly preferred means for accomplishing this are by exposing the polypeptide to glycosylating enzymes derived from cells which normally provide such processing, e.g., mammalian glycosylation enzymes. Deglycosylation enzymes are also contemplated. Also embraced are versions of the same primary amino acid sequence which have other minor modifications, including phosphorylated amino acid residues, e.g., phosphotyrosine, phosphoserine, or phosphothreonine, or nucleoside or nucleotide derivatives, e.g., guanyl derivatized.

A major group of derivatives are covalent conjugates of the respective chemokine or receptor or fragments thereof with other proteins or polypeptides. These derivatives can be synthesized in recombinant culture such as N- or C-terminal fusions or by the use of agents known in the art for their usefulness in cross-linking proteins through reactive side groups. Preferred chemokine derivatization sites with cross-linking agents are at free amino groups, carbohydrate moieties, and cysteine residues.

Fusion polypeptides between these chemokines or receptors and other homologous or heterologous proteins, e.g., other chemokines or receptors, are also provided. Many growth factors

and cytokines are homodimeric entities, and a repeat construct may have various advantages, including lessened susceptibility to proteolytic cleavage. Moreover, many cytokine receptors require dimerization to transduce a signal, and various dimeric ligands or domain repeats can be desirable. Homologous polypeptides may be fusions between different surface markers, resulting in, e.g., a hybrid protein exhibiting receptor binding specificity. Likewise, heterologous fusions may be constructed which would exhibit a combination of properties or activities of the derivative proteins. Typical examples are fusions of a reporter polypeptide, e.g., luciferase, with a segment or domain of a ligand, e.g., a receptor-binding segment, so that the presence or location of the fused ligand, or a binding composition, may be easily determined. See, e.g., Dull, et al., U.S. Patent No. 4,859,609. Other gene fusion partners include bacterial ß-galactosidase, trpE, Protein A, ß-lactamase, alpha amylase, alcohol dehydrogenase, a FLAG fusion, and yeast alpha mating factor. See, e.g., Godowski, et al. (1988) Science 241:812-816.

The phosphoramidite method described by Beaucage and Carruthers (1981) Tetra. Letts. 22:1859-1862, will produce suitable synthetic DNA fragments. A double stranded fragment will often be obtained either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

Such polypeptides may also have amino acid residues which have been chemically modified by phosphorylation, guanylation, sulfonation, biotinylation, or the addition or removal of other moieties, particularly those which have molecular shapes similar to phosphate or guanyl groups. In some embodiments, the modifications will be useful labeling reagents, or serve as purification targets, e.g., affinity tags as FLAG.

Fusion proteins will typically be made by either recombinant nucleic acid methods or by synthetic polypeptide methods. Techniques for nucleic acid manipulation and expression are described generally, for example, in Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed.), Vols. 1-3, Cold Spring

Harbor Laboratory. Techniques for synthesis of polypeptides are described, for example, in Merrifield (1963) J. Amer. Chem. Soc. 85:2149-2156; Merrifield (1986) Science 232: 341-347; and Atherton, et al. (1989) Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford; and chemical ligation, e.g., Dawson, et al. (1994) Science 266:776-779, a method of linking long synthetic peptides by a peptide bond.

This invention also contemplates the use of derivatives of these chemokines or receptors other than variations in amino acid sequence or glycosylation. Such derivatives may involve covalent or aggregative association with chemical moieties. These derivatives generally include: (1) salts, (2) side chain and terminal residue covalent modifications, and (3) adsorption complexes, for example with cell membranes. Such covalent or aggregative derivatives are useful as immunogens, as reagents in immunoassays, or in purification methods such as for affinity purification of ligands or other binding ligands. For example, a chemokine antigen can be immobilized by covalent bonding to a solid support such as cyanogen bromide-activated Sepharose, by methods which are well known in the art, or adsorbed onto polyolefin surfaces, with or without glutaraldehyde cross-linking, for use in the assay or purification of anti-chemokine antibodies or its receptor. These chemokines can also be labeled with a detectable group, for example radioiodinated by the chloramine T procedure, covalently bound to rare earth chelates, or conjugated to a fluorescent moiety for use in diagnostic assays. Purification of chemokine, receptor, or binding compositions may be effected by immobilized antibodies or receptor.

Other modifications may be introduced with the goal of modifying the therapeutic pharmacokinetics or pharmacodynamics of a target chemokine. For example, certain means to minimize the size of the entity may improve its pharmacoaccessibility; other means to maximize size may affect pharmacodynamics. Similarly, changes in ligand binding kinetics or equilibrium of a receptor may be engineered.

A solubilized chemokine or receptor or appropriate fragment of this invention can be used as an immunogen for the production of antisera or antibodies specific for the ligand, receptor, or fragments thereof. The purified proteins can be used to screen monoclonal antibodies or chemokine-binding fragments prepared by immunization with various forms of impure preparations containing the protein. In particular, antibody equivalents include antigen binding fragments of natural antibodies, e.g., Fv, Fab, or $F(ab)_2$. Purified chemokines can also be used as a reagent to detect antibodies generated in response to the presence of elevated levels of the protein or cell fragments containing the protein, both of which may be diagnostic of an abnormal or specific physiological or disease condition. Additionally, chemokine protein fragments, or their concatenates, may also serve as immunogens to produce binding compositions, e.g., antibodies of the present invention, as described immediately below. For example, this invention contemplates antibodies raised against certain amino acid sequences, e.g., in in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 and 26, or proteins containing them. In particular, this invention contemplates antibodies having binding affinity to or being raised against specific fragments, e.g., those which are predicted to lie on the outside surfaces of protein tertiary structure. Similar concepts apply to antibodies specific for receptors of the invention.

The present invention contemplates the isolation of additional closely related species variants. Southern and Northern blot analysis should establish that similar genetic entities exist in other related mammals, and establish the stringency of hybridization conditions to isolate such. It is likely that these chemokines and receptors are widespread in species variants, e.g., among the rodents and the primates.

The invention also provides means to isolate a group of related chemokines or receptors displaying both distinctness and similarities in structure, expression, and function. Elucidation of many of the physiological effects of the proteins will be greatly accelerated by the isolation and characterization of distinct species variants of the ligands. Related genes found, e.g., in various

computer databases will also be useful, in many instances, for similar purposes with structurally related proteins. In particular, the present invention provides useful probes or search features for identifying additional homologous genetic entities in different species.

The isolated genes will allow transformation of cells lacking expression of a corresponding chemokine or receptor, e.g., either species types or cells which lack corresponding antigens and exhibit negative background activity. Expression of transformed genes will allow isolation of antigenically pure cell lines, with defined or single specie variants. This approach will allow for more sensitive detection and discrimination of the physiological effects of chemokine or receptor proteins. Subcellular fragments, e.g., cytoplasts or membrane fragments, can be isolated and used.

Dissection of critical structural elements which effect the various differentiation functions provided by ligands is possible using standard techniques of modern molecular biology, particularly in comparing members of the related class. See, e.g., the homolog-scanning mutagenesis technique described in Cunningham, et al. (1989) Science 243:1339-1336; and approaches used in O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992; and Lechleiter, et al. (1990) EMBO J. 9:4381-4390.

In addition, various segments can be substituted between species variants to determine what structural features are important in both receptor binding affinity and specificity, as well as signal transduction. An array of different chemokine or receptor variants will be used to screen for variants exhibiting combined properties of interaction with different species variants.

Intracellular functions would probably involve segments of the receptor which are normally accessible to the cytosol. However, ligand internalization may occur under certain circumstances, and interaction between intracellular components and "extracellular" segments may occur. The specific segments of interaction of a particular chemokine with other intracellular components may be identified by mutagenesis or direct biochemical means, e.g., cross-linking or affinity methods. Structural analysis by crystallographic

or other physical methods will also be applicable. Further investigation of the mechanism of signal transduction will include study of associated components which may be isolatable by affinity methods or by genetic means, e.g., complementation analysis of mutants.

Further study of the expression and control of the various chemokines or receptors will be pursued. The controlling elements associated with the proteins may exhibit differential developmental, tissue specific, or other expression patterns. Upstream or downstream genetic regions, e.g., control elements, are of interest. Differential splicing of message may lead to membrane bound forms, soluble forms, and modified versions of ligand.

Structural studies of the proteins will lead to design of new ligands or receptors, particularly analogs exhibiting agonist or antagonist properties on the receptor. This can be combined with previously described screening methods to isolate ligands exhibiting desired spectra of activities.

Expression in other cell types will often result in glycosylation differences in a particular chemokine or receptor. Various species variants may exhibit distinct functions based upon structural differences other than amino acid sequence. Differential modifications may be responsible for differential function, and elucidation of the effects are now made possible.

Thus, the present invention provides important reagents related to a physiological ligand-receptor interaction. Although the foregoing description has focused primarily upon the mouse and human embodiments of the chemokines or receptors specifically described, those of skill in the art will immediately recognize that the invention provides other counterparts, e.g., from related species, rodents or primates.

V. Antibodies

Antibodies can be raised to these chemokines or receptors, including species or polymorphic variants, and fragments thereof, both in their naturally occurring forms and in their recombinant forms. Additionally, antibodies can be raised to chemokines or

receptors in either their active or inactive forms, or in their native or denatured forms. Anti-idiotypic antibodies are also contemplated.

Antibodies, including binding fragments and single chain versions, against predetermined fragments of the ligands can be raised by immunization of animals with concatemers or conjugates of the fragments with immunogenic proteins. Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies can be screened for binding to normal or defective chemokines or receptors, or screened for agonistic or antagonistic activity. These monoclonal antibodies will usually bind with at least a K_D of about 1 mM, more usually at least about 300 μM, typically at least about 10 μM, more typically at least about 30 μM, preferably at least about 10 μM, and more preferably at least about 3 μM or better.

The antibodies, including antigen binding fragments, of this invention can have significant preparative, diagnostic, or therapeutic value. They can be useful to purify or label the desired antigen in a sample, or may be potent antagonists that bind to ligand and inhibit binding to receptor or inhibit the ability of a ligand to elicit a biological response. They also can be useful as non-neutralizing antibodies and can be coupled to, or as fusion proteins with, toxins or radionuclides so that when the antibody binds to antigen, a cell expressing it, e.g., on its surface via receptor, is killed. Further, these antibodies can be conjugated to drugs or other therapeutic agents, either directly or indirectly by means of a linker, and may effect drug targeting. Antibodies to receptors may be more easily used to block ligand binding and/or signal transduction.

The antibodies of this invention can also be useful in diagnostic or reagent purification applications. As capture or non-neutralizing antibodies, they can be screened for ability to bind to the chemokines or receptors without inhibiting ligand-receptor binding. As neutralizing antibodies, they can be useful in competitive binding assays. They will also be useful in detecting or quantifying chemokine or receptors, e.g., in immunoassays. They may be used as purification reagents in immunoaffinity columns or as immunohistochemistry reagents.

Ligand or receptor fragments may be concatenated or joined to other materials, particularly polypeptides, as fused or covalently joined polypeptides to be used as immunogens. Short peptides will preferably be made as repeat structures to increase size. A ligand and its fragments may be fused or covalently linked to a variety of immunogens, such as keyhole limpet hemocyanin, bovine serum albumin, tetanus toxoid, etc. See Microbiology, Hoeber Medical Division, Harper and Row, 1969; Landsteiner (1962) Specificity of Serological Reactions, Dover Publications, New York, and Williams, et al. (1967) Methods in Immunology and Immunochemistry, Vol. 1, Academic Press, New York, for descriptions of methods of preparing polyclonal antisera. A typical method involves hyperimmunization of an animal with an antigen. The blood of the animal is then collected shortly after the repeated immunizations and the gamma globulin fraction is isolated.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description of techniques for preparing such monoclonal antibodies may be found in, e.g., Stites, et al. (eds.) Basic and Clinical Immunology (4th ed.), Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH Press; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press, New York; and particularly in Kohler and Milstein (1975) in Nature 256:495-497, which discusses one method of generating monoclonal antibodies. Summarized briefly, this method involves injecting an animal with an immunogen. The animal is then sacrificed and cells taken, e.g., from its spleen, which are then fused with myeloma cells. The result is a hybrid cell or "hybridoma" that is capable of reproducing in vitro. The population of hybridomas is then screened to isolate individual clones, each of which secrete a single antibody species to the immunogen. In this manner, the individual antibody species obtained are the products of immortalized and cloned single B cells from the immune animal generated in response to a specific site recognized on the

immunogenic substance. Large amounts of antibody may be derived from ascites fluid from an animal.

Other suitable techniques involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors. See, Huse, et al. (1989) "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda," Science 246:1275-1281; and Ward, et al. (1989) Nature 341:544-546. The polypeptides and antibodies of the present invention may be used with or without modification, including chimeric or humanized antibodies. Frequently, the polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents, teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced, see Cabilly, U.S. Patent No. 4,816,567; and Queen et al. (1989) Proc. Nat'l. Acad. Sci. 86:10029-10033.

The antibodies of this invention can also be used for affinity chromatography in isolating the protein. Columns can be prepared where the antibodies are linked to a solid support, e.g., particles, such as agarose, Sephadex, or the like, where a cell lysate may be passed through the column, the column washed, followed by increasing concentrations of a mild denaturant, whereby the purified chemokine protein will be released.

The antibodies may also be used to screen expression libraries for particular expression products. Usually the antibodies used in such a procedure will be labeled with a moiety allowing easy detection of presence of antigen by antibody binding.

Antibodies raised against these chemokines or receptors will also be useful to raise anti-idiotypic antibodies. These will be useful

in detecting or diagnosing various immunological conditions related to expression of the respective antigens.

VI. Nucleic Acids

The described peptide sequences and the related reagents are useful in isolating a DNA clone encoding these chemokines or receptors, e.g., from a natural source. Typically, it will be useful in isolating a gene from another individual, and similar procedures will be applied to isolate genes from related species, e.g., rodents or primates. Cross hybridization will allow isolation of ligand from other closely related species. A number of different approaches should be available to successfully isolate a suitable nucleic acid clone. Similar concepts apply to the receptor embodiments.

The purified protein or defined peptides are useful for generating antibodies by standard methods, as described above. Synthetic peptides or purified protein can be presented to an immune system to generate monoclonal or polyclonal antibodies. See, e.g., Coligan (1991) Current Protocols in Immunology Wiley/Greene; and Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press. Alternatively, a chemokine or receptor may be used as a specific binding reagent, and advantage can be taken of its specificity of binding, much like an antibody would be used. The chemokine receptors are typically 7 transmembrane proteins, which could be sensitive to appropriate interaction with lipid or membrane. The signal transduction typically is mediated through a G-protein, through interaction with a G-protein coupled receptor.

For example, the specific binding composition could be used for screening of an expression library made from a cell line which expresses a particular chemokine. The screening can be standard staining of surface expressed ligand, or by panning. Screening of intracellular expression can also be performed by various staining or immunofluorescence procedures. The binding compositions could be used to affinity purify or sort out cells expressing the ligand.

The peptide segments can also be used to predict appropriate oligonucleotides to screen a library, e.g., to isolate species variants.

The genetic code can be used to select appropriate oligonucleotides useful as probes for screening. n combination with polymerase chain reaction (PCR) techniques, synthetic oligonucleotides will be useful in selecting correct clones from a library. Complementary sequences will also be used as probes or primers. Anchored vector or poly-A complementary PCR techniques or complementary DNA of other peptides may be useful.

This invention contemplates use of isolated DNA or fragments to encode a biologically active corresponding chemokine polypeptide. In addition, this invention covers isolated or recombinant DNA which encodes a biologically active protein or polypeptide which is capable of hybridizing under appropriate conditions with the DNA sequences described herein. Said biologically active protein or polypeptide can be an intact ligand. receptor, or fragment, and have an amino acid sequence as disclosed in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 or 26. Further, this invention covers the use of isolated or recombinant DNA, or fragments thereof, which encode proteins which are homologous to a chemokine or receptor or which was isolated using such a cDNA encoding a chemokine or receptor as a probe. The isolated DNA can have the respective regulatory sequences in the 5' and 3' flanks, e.g., promoters, enhancers, poly-A addition signals, and others.

An "isolated" nucleic acid is a nucleic acid, e.g., an RNA, DNA, or a mixed polymer, which is substantially separated from other components which naturally accompany a native sequence, e.g., ribosomes, polymerases, and flanking genomic sequences from the originating species. The term embraces a nucleic acid sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs or analogs biologically synthesized by heterologous systems. A substantially pure molecule includes isolated forms of the molecule.

An isolated nucleic acid will generally be a homogeneous composition of molecules, but will, in some embodiments, contain minor heterogeneity. This heterogeneity is typically found at the

polymer ends or portions not critical to a desired biological function or activity.

A "recombinant" nucleic acid is defined either by its method of production or its structure. In reference to its method of production, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques, e.g., involving human intervention in the nucleotide sequence, typically selection or production. Alternatively, it can be a nucleic acid made by generating a sequence comprising fusion of two fragments which are not naturally contiguous to each other, but is meant to exclude products of nature, e.g., naturally occurring purified forms. Thus, for example, products made by transforming cells with any unnaturally occurring vector is encompassed, as are nucleic acids comprising sequence derived using a synthetic oligonucleotide process. Such is often done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a single genetic entity comprising a desired combination of functions not found in the commonly available natural forms. Restriction enzyme recognition sites are often the target of such artificial manipulations, but other site specific targets, e.g., promoters, DNA replication sites, regulation sequences, control sequences, or other useful features may be incorporated by design. A similar concept is intended for a recombinant, e.g., fusion, polypeptide. Specifically included are synthetic nucleic acids which, by genetic code redundancy, encode polypeptides similar to fragments of these antigens, and fusions of sequences from various different species variants.

A significant "fragment" in a nucleic acid context is a contiguous segment of at least about 17 nucleotides, generally at least about 20 nucleotides, more generally at least about 23 nucleotides, ordinarily at least about 26 nucleotides, more ordinarily at least about 29 nucleotides, often at least about 32 nucleotides, more often at least about 35 nucleotides, typically at least about 38 nucleotides, more typically at least about 41 nucleotides, usually at least about 44

nucleotides, more usually at least about 47 nucleotides, preferably at least about 50 nucleotides, more preferably at least about 53 nucleotides, and in particularly preferred embodiments will be at least about 56 or more nucleotides, e.g., 60, 65, 75, 85, 100, 120, 150, 200, 250, 300, 400, etc. Such fragments may have ends which begin and/or end at virtually all positions, e.g., beginning at nucleotides 1, 2, 3, etc., and ending at, e.g., 300, 299, 298, 287, etc., in combinatorial pairs. Particularly interesting polynucleotides have ends corresponding to structural domain boundaries.

A DNA which codes for a particular chemokine or receptor protein or peptide will be very useful to identify genes, mRNA, and cDNA species which code for related or homologous ligands or receptors, as well as DNAs which code for homologous proteins from different species. There are likely homologs in closely related species, e.g., rodents or primates. Various chemokine proteins should be homologous and are encompassed herein, as would be receptors. However, proteins can readily be isolated under appropriate conditions using these sequences if they are sufficiently homologous. Typically, primate chemokines or receptors are of particular interest.

This invention further covers recombinant DNA molecules and fragments having a DNA sequence identical to or highly homologous to the isolated DNAs set forth herein. In particular, the sequences will often be operably linked to DNA segments which control transcription, translation, and DNA replication. Alternatively, recombinant clones derived from the genomic sequences, e.g., containing introns, will be useful for transgenic studies, including, e.g., transgenic cells and organisms, and for gene therapy. See, e.g., Goodnow (1992) "Transgenic Animals" in Roitt (ed.) Encyclopedia of Immunology Academic Press, San Diego, pp. 1502-1504; Travis (1992) Science 256:1392-1394; Kuhn, et al. (1991) Science 254:707-710; Capecchi (1989) Science 244:1288; Robertson (1987)(ed.) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach IRL Press, Oxford; and Rosenberg (1992) J. Clinical Oncology 10:180-199.

Homologous nucleic acid sequences, when compared, exhibit significant similarity, or identity. The standards for homology in nucleic acids are either measures for homology generally used in the art by sequence comparison or based upon hybridization conditions. The hybridization conditions are described in greater detail below.

Substantial homology in the nucleic acid sequence comparison context means either that the segments, or their complementary strands, when compared, are identical when optimally aligned, with appropriate nucleotide insertions or deletions, in at least about 50% of the nucleotides, generally at least about 56%, more generally at least about 59%, ordinarily at least about 62%, more ordinarily at least about 65%, often at least about 68%, more often at least about 71%, typically at least about 74%, more typically at least about 77%, usually at least about 80%, more usually at least about 85%, preferably at least about 90%, more preferably at least about 95 to 98% or more, and in particular embodiments, as high at about 99% or more of the nucleotides. Alternatively, substantial homology exists when the segments will hybridize under selective hybridization conditions, to a strand, or its complement, typically using a sequence derived from SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23 or 25. Typically, selective hybridization will occur when there is at least about 55% homology over a stretch of at least about 30 nucleotides, preferably at least about 65% over a stretch of at least about 25 nucleotides, more preferably at least about 75%, and most preferably at least about 90% over about 20 nucleotides. See, Kanehisa (1984) Nuc. Acids Res. 12:203-213. The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will be over a stretch of at least about 17 nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 40 nucleotides, preferably at least about 50 nucleotides, and more preferably at least about 75 to 100 or more nucleotides. PCR primers will generally have high levels of matches over potentially shorter lengths.

Stringent conditions, in referring to homology in the hybridization context, will be stringent combined conditions of salt,

temperature, organic solvents, and other parameters, typically those controlled in hybridization reactions. Stringent temperature conditions will usually include temperatures in excess of about 30° C, more usually in excess of about 37° C, typically in excess of about 45° C, more typically in excess of about 55° C, preferably in excess of about 65° C, and more preferably in excess of about 70° C. Stringent salt conditions will ordinarily be less than about 1000 mM, usually less than about 500 mM, more usually less than about 400 mM, typically less than about 300 mM, preferably less than about 200 mM, and more preferably less than about 150 mM, e.g., 20-50 mM. However, the combination of parameters is much more important than the measure of any single parameter. See, e.g., Wetmur and Davidson (1968) J. Mol. Biol. 31:349-370.

Corresponding chemokines or receptors from other closely related species can be cloned and isolated by cross-species hybridization. Alternatively, sequences from a sequence data base may be recognized as having similarity. Homology may be very low between distantly related species, and thus hybridization of relatively closely related species is advisable. Alternatively, preparation of an antibody preparation which exhibits less species specificity may be useful in expression cloning approaches. PCR approaches using segments of conserved sequences will also be used.

VII. Making Chemokines or Receptors; Mimetics

DNA which encodes each respective chemokine, receptor, or fragments thereof can be obtained by chemical synthesis, screening cDNA libraries, or by screening genomic libraries prepared from a wide variety of cell lines or tissue samples.

This DNA can be expressed in a wide variety of host cells for the synthesis of a full-length ligand or fragments which can in turn, for example, be used to generate polyclonal or monoclonal antibodies; for binding studies; for construction and expression of modified molecules; for expression cloning or purification; and for structure/function studies. Each antigen or its fragments can be expressed in host cells that are transformed or transfected with appropriate expression vectors. These molecules can be substantially

purified to be free of protein or cellular contaminants, other than those derived from the recombinant host, and therefore are particularly useful in pharmaceutical compositions when combined with a pharmaceutically acceptable carrier and/or diluent. The antigens or antibodies, or portions thereof, may be expressed as fusions with other proteins.

Expression vectors are typically self-replicating DNA or RNA constructs containing the desired antigen gene or its fragments, usually operably linked to suitable genetic control elements that are recognized in a suitable host cell. These control elements are capable of effecting expression within a suitable host. The specific type of control elements necessary to effect expression will depend upon the eventual host cell used. Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system, and typically include a transcriptional promoter, an optional operator to control the onset of transcription, transcription enhancers to elevate the level of mRNA expression, a sequence that encodes a suitable ribosome binding site, and sequences that terminate transcription and translation. Expression vectors also usually contain an origin of replication that allows the vector to replicate independently of the host cell.

The vectors of this invention contain DNA which encode embodiments of a chemokine, receptor, or a fragment thereof, typically encoding a biologically active polypeptide. The DNA can be under the control of a viral promoter and can encode a selection marker. This invention further contemplates use of such expression vectors which are capable of expressing eukaryotic cDNA coding for each chemokine or receptor in a prokaryotic or eukaryotic host, where the vector is compatible with the host and where the eukaryotic cDNA coding for the protein is inserted into the vector such that growth of the host containing the vector expresses the cDNA in question. Usually, expression vectors are designed for stable replication in their host cells or for amplification to greatly increase the total number of copies of the desirable gene per cell. It is not always necessary to require that an expression vector replicate in a host cell, e.g., it is possible to effect transient expression of the

ligand or its fragments in various hosts using vectors that do not contain a replication origin that is recognized by the host cell. It is also possible to use vectors that cause integration of a chemokine or receptor gene or its fragments into the host DNA by recombination, or to integrate a promoter which controls expression of an endogenous gene.

Vectors, as used herein, comprise plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles, including those which enable the integration of DNA fragments into the genome of the host. Expression vectors are specialized vectors which contain genetic control elements that effect expression of operably linked genes. Plasmids are the most commonly used form of vector but many other forms of vectors which serve an equivalent function and which are, or become, known in the art are suitable for use herein. See, e.g., Pouwels, et al. (1985 and Supplements) Cloning Vectors: A Laboratory Manual, Elsevier, N.Y., and Rodriquez, et al. (1988)(eds.) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Buttersworth, Boston, MA.

Transformed cells include cells, preferably mammalian, that have been transformed or transfected with a chemokine or receptor gene containing vector constructed using recombinant DNA techniques. Transformed host cells usually express the ligand, receptor, or its fragments, but for purposes of cloning, amplifying, and manipulating its DNA, do not need to express the protein. This invention further contemplates culturing transformed cells in a nutrient medium, thus permitting the protein to accumulate in the culture. The protein can be recovered, from the culture or from the culture medium, or from cell membranes.

For purposes of this invention, DNA sequences are operably linked when they are functionally related to each other. For example, DNA for a presequence or secretory signal is operably linked to a polypeptide if it is expressed as a preprotein or participates in directing the polypeptide to the cell membrane or in secretion of the polypeptide. A promoter is operably linked to a coding sequence if it controls the transcription of the polypeptide; a ribosome binding site is operably linked to a coding sequence if it is

positioned to permit translation. Usually, operably linked means contiguous and in reading frame, however, certain genetic elements such as repressor genes are not contiguously linked but still bind to operator sequences that in turn control expression.

Suitable host cells include prokaryotes, lower eukaryotes, and higher eukaryotes. Prokaryotes include both gram negative and gram positive organisms, e.g., E. coli and B. subtilis. Lower eukaryotes include yeasts, e.g., S. cerevisiae and Pichia, and species of the genus Dictyostelium. Higher eukaryotes include established tissue culture cell lines from animal cells, both of non-mammalian origin, e.g., insect cells, and birds, and of mammalian origin, e.g., human, primates, and rodents.

Prokaryotic host-vector systems include a wide variety of vectors for many different species. As used herein, E. coli and its vectors will be used generically to include equivalent vectors used in other prokaryotes. A representative vector for amplifying DNA is pBR322 or many of its derivatives. Vectors that can be used to express these chemokines or their fragments include, but are not limited to, such vectors as those containing the lac promoter (pUC-series); trp promoter (pBR322-trp); Ipp promoter (the pIN-series); lambda-pP or pR promoters (pOTS); or hybrid promoters such as ptac (pDR540). See Brosius, et al. (1988) "Expression Vectors Employing Lambda-, trp-, lac-, and Ipp-derived Promoters", in Rodriguez and Denhardt (eds.) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Buttersworth, Boston, Chapter 10, pp. 205-236.

Lower eukaryotes, e.g., yeasts and Dictyostelium, may be transformed with chemokine or receptor sequence containing nucleic acids. For purposes of this invention, the most common lower eukaryotic host is the baker's yeast, Saccharomyces cerevisiae. It will be used to generically represent lower eukaryotes although a number of other strains and species are also available. Yeast vectors typically consist of a replication origin (unless of the integrating type), a selection gene, a promoter, DNA encoding the desired protein or its fragments, and sequences for translation termination, polyadenylation, and transcription termination. Suitable expression vectors for yeast include such constitutive promoters as 3-

phosphoglycerate kinase and various other glycolytic enzyme gene promoters or such inducible promoters as the alcohol dehydrogenase 2 promoter or metallothionine promoter. Suitable vectors include derivatives of the following types: self-replicating low copy number (such as the YRp-series), self-replicating high copy number (such as the YEp-series); integrating types (such as the YIp-series), or mini-chromosomes (such as the YCp-series).

Higher eukaryotic tissue culture cells are the preferred host cells for expression of the functionally active chemokine or receptor proteins. In principle, most any higher eukaryotic tissue culture cell line is workable, e.g., insect baculovirus expression systems, whether from an invertebrate or vertebrate source. However, mammalian cells are preferred, in that the processing, both cotranslationally and posttranslationally, will be typically most like natural. Transformation or transfection and propagation of such cells has become a routine procedure. Examples of useful cell lines include HeLa cells, Chinese hamster ovary (CHO) cell lines, baby rat kidney (BRK) cell lines, insect cell lines, bird cell lines, and monkey (COS) cell lines. Expression vectors for such cell lines usually include an origin of replication, a promoter, a translation initiation site, RNA splice sites (if genomic DNA is used), a polyadenylation site, and a transcription termination site. These vectors also usually contain a selection gene or amplification gene. Suitable expression vectors may be plasmids, viruses, or retroviruses carrying promoters derived, e.g., from such sources as from adenovirus, SV40, parvoviruses, vaccinia virus, or cytomegalovirus. Representative examples of suitable expression vectors include pCDNA1; pCD, see Okayama, et al. (1985) Mol. Cell Biol. 5:1136-1142; pMC1neo Poly-A, see Thomas, et al. (1987) Cell 51:503-512; and a baculovirus vector such as pAC 373 or pAC 610.

It will often be desired to express a chemokine or receptor polypeptide in a system which provides a specific or defined glycosylation pattern. In this case, the usual pattern will be that provided naturally by the expression system. However, the pattern will be modifiable by exposing the polypeptide, e.g., an unglycosylated form, to appropriate glycosylating proteins

introduced into a heterologous expression system. For example, a chemokine or receptor gene may be co-transformed with one or more genes encoding mammalian or other glycosylating enzymes. Using this approach, certain mammalian glycosylation patterns will be achievable or approximated in prokaryote or other cells.

A chemokine, receptor, or a fragment thereof, may be engineered to be phosphatidyl inositol (PI) linked to a cell membrane, but can be removed from membranes by treatment with a phosphatidyl inositol cleaving enzyme, e.g., phosphatidyl inositol phospholipase-C. This releases the antigen in a biologically active form, and allows purification by standard procedures of protein chemistry. See, e.g., Low (1989) Biochim. Biophys. Acta 988:427-454; Tse, et al. (1985) Science 230:1003-1008; and Brunner, et al. (1991) J. Cell Biol. 114:1275-1283.

Now that these chemokines and receptors have been characterized, fragments or derivatives thereof can be prepared by conventional processes for synthesizing peptides. These include processes such as are described in Stewart and Young (1984) Solid Phase Peptide Synthesis, Pierce Chemical Co., Rockford, IL; Bodanszky and Bodanszky (1984) The Practice of Peptide Synthesis, Springer-Verlag, New York; and Bodanszky (1984) The Principles of Peptide Synthesis, Springer-Verlag, New York. For example, an azide process, an acid chloride process, an acid anhydride process, a mixed anhydride process, an active ester process (for example, p-nitrophenyl ester, N-hydroxysuccinimide ester, or cyanomethyl ester), a carbodiimidazole process, an oxidative-reductive process, or a dicyclohexyl-carbodiimide (DCCD)/additive process can be used. Solid phase and solution phase syntheses are both applicable to the foregoing processes.

These chemokines, receptors, fragments, or derivatives are suitably prepared in accordance with the above processes as typically employed in peptide synthesis, generally either by a so-called stepwise process which comprises condensing an amino acid to the terminal amino acid, one by one in sequence, or by coupling peptide fragments to the terminal amino acid. Amino groups that are not

being used in the coupling reaction are typically protected to prevent coupling at an incorrect location.

If a solid phase synthesis is adopted, the C-terminal amino acid is typically bound to an insoluble carrier or support through its carboxyl group. The insoluble carrier is not particularly limited as long as it has a binding capability to a reactive carboxyl group. Examples of such insoluble carriers include halomethyl resins, such as chloromethyl resin or bromomethyl resin, hydroxymethyl resins, phenol resins, tert-alkyloxycarbonyl-hydrazidated resins, and the like.

An amino group-protected amino acid is bound in sequence through condensation of its activated carboxyl group and the reactive amino group of the previously formed peptide or chain, to synthesize the peptide step by step. After synthesizing the complete sequence, the peptide is split off from the insoluble carrier to produce the peptide. This solid-phase approach is generally described, e.g., by Merrifield, et al. (1963) in J. Am. Chem. Soc. 85:2149-2156.

The prepared ligand and fragments thereof can be isolated and purified from the reaction mixture by means of peptide separation, e.g., by extraction, precipitation, electrophoresis, and various forms of chromatography, and the like. The various chemokines or receptors of this invention can be obtained in varying degrees of purity depending upon its desired use. Purification can be accomplished by use of the protein purification techniques disclosed herein or by the use of the antibodies herein described, e.g., in immunoabsorbant affinity chromatography. This immunoabsorbant affinity chromatography is typically carried out, e.g., by first linking the antibodies to a solid support and then contacting the linked antibodies with solubilized lysates of appropriate source cells, lysates of other cells expressing the ligand or receptor, or lysates or supernatants of cells producing the desired proteins as a result of DNA techniques, see below.

VIII. Uses

The present invention provides reagents which will find use in diagnostic applications as described elsewhere herein, e.g., in the general description for developmental abnormalities, or below in the description of kits for diagnosis.

This invention also provides reagents with significant therapeutic potential. These chemokines and receptors (naturally occurring or recombinant), fragments thereof, and binding compositions, e.g., antibodies thereto, along with compounds identified as having binding affinity to them, should be useful in the treatment of conditions associated with abnormal physiology or development, including inflammatory conditions, e.g., asthma. In particular, modulation of trafficking of leukocytes is one likely biological activity, but a wider tissue distribution might suggest broader biological activity, including, e.g., antiviral effects. Abnormal proliferation, regeneration, degeneration, and atrophy may be modulated by appropriate therapeutic treatment using the compositions provided herein. For example, a disease or disorder associated with abnormal expression or abnormal signaling by a chemokine or ligand for a receptor should be a likely target for an agonist or antagonist of the ligand.

Various abnormal physiological or developmental conditions are known in cell types shown to possess the chemokine or receptor mRNAs by Northern blot analysis. See Berkow (ed.) The Merck Manual of Diagnosis and Therapy, Merck & Co., Rahway, N.J.; and Thorn, et al. Harrison's Principles of Internal Medicine, McGraw-Hill, N.Y. Developmental or functional abnormalities, e.g., of the immune system, cause significant medical abnormalities and conditions which may be susceptible to prevention or treatment using compositions provided herein.

Antibodies to the chemokines or receptors, including recombinant forms, can be purified and then used diagnostically or therapeutically, alone or in combination with other chemokines, cytokines, or antagonists thereof. These reagents can be combined for therapeutic use with additional active or inert ingredients, e.g., in conventional pharmaceutically acceptable carriers or diluents, e.g.,

immunogenic adjuvants, along with physiologically innocuous stabilizers and excipients. These combinations can be sterile filtered and placed into dosage forms as by lyophilization in dosage vials or storage in stabilized aqueous preparations. This invention also contemplates use of antibodies or binding fragments thereof, including forms which are not complement binding. Moreover, modifications to the antibody molecules or antigen binding fragments thereof, may be adopted which affect the pharmacokinetics or pharmacodynamics of the therapeutic entity.

Drug screening using antibodies or receptor or fragments thereof can be performed to identify compounds having binding affinity to each chemokine or receptor, including isolation of associated components. Subsequent biological assays can then be utilized to determine if the compound has intrinsic stimulating activity and is therefore a blocker or antagonist in that it blocks the activity of the ligand. Likewise, a compound having intrinsic stimulating activity can activate the receptor and is thus an agonist in that it simulates the activity of a ligand. This invention further contemplates the therapeutic use of antibodies to these chemokines as antagonists, or to the receptors as antagonists or agonists. This approach should be particularly useful with other chemokine or receptor species variants.

The quantities of reagents necessary for effective therapy will depend upon many different factors, including means of administration, target site, physiological state of the patient, and other medicants administered. Thus, treatment dosages should be titrated to optimize safety and efficacy in various populations, including racial subgroups, age, gender, etc. Typically, dosages used in vitro may provide useful guidance in the amounts useful for in situ administration of these reagents. Animal testing of effective doses for treatment of particular disorders will provide further predictive indication of human dosage. Various considerations are described, e.g., in Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics, 8th Ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Penn.. Methods for administration are

discussed therein and below, e.g., for oral, intravenous, intraperitoneal, or intramuscular administration, transdermal diffusion, and others. Pharmaceutically acceptable carriers typically include water, saline, buffers, and other compounds described, e.g., in the Merck Index, Merck & Co., Rahway, New Jersey. Dosage ranges would ordinarily be expected to be in amounts lower than 1 mM concentrations, typically less than about 10 μM concentrations, usually less than about 100 nM, preferably less than about 10 pM (picomolar), and most preferably less than about 1 fM (femtomolar), with an appropriate carrier. Slow release formulations, or a slow release apparatus will often be utilized for continuous administration.

A chemokine, fragments thereof, or antibodies to it or its fragments, antagonists, and agonists, may be administered directly to the host to be treated or, depending on the size of the compounds, it may be desirable to conjugate them to carrier proteins such as ovalbumin or serum albumin prior to their administration. Therapeutic formulations may be administered in many conventional dosage formulations. While it is possible for the active ingredient to be administered alone, it is often preferable to present it as a pharmaceutical formulation. Formulations typically comprise at least one active ingredient, as defined above, together with one or more acceptable carriers thereof. Each carrier should be both pharmaceutically and physiologically acceptable in the sense of being compatible with the other ingredients and not injurious to the patient. Carriers may improve storage life, stability, etc. Formulations include those suitable for oral, rectal, nasal, or parenteral (including subcutaneous, intramuscular, intravenous and intradermal) administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well known in the art of pharmacy. See, e.g., Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics, 8th Ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Penn.; Avis, et al. (eds.) (1993) Pharmaceutical Dosage Forms: Parenteral Medications Dekker, New York; Lieberman, et al. (eds.)

(1990) Pharmaceutical Dosage Forms: Tablets Dekker, New York; and Lieberman, et al. (eds.) (1990) Pharmaceutical Dosage Forms: Disperse Systems Dekker, New York. The therapy of this invention may be combined with or used in association with other therapeutic agents. Similar considerations will often apply to receptor based reagents.

Both the naturally occurring and the recombinant forms of the chemokines or receptors of this invention are particularly useful in kits and assay methods which are capable of screening compounds for binding activity to the proteins. Several methods of automating assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period. See, e.g., Fodor, et al. (1991) Science 251:767-773, which describes means for testing of binding affinity by a plurality of defined polymers synthesized on a solid substrate. The development of suitable assays can be greatly facilitated by the availability of large amounts of purified, soluble chemokine as provided by this invention.

For example, antagonists can normally be found once a ligand has been structurally defined. Testing of potential ligand analogs is now possible upon the development of highly automated assay methods using physiologically responsive cells. In particular, new agonists and antagonists will be discovered by using screening techniques described herein.

Viable cells could also be used to screen for the effects of drugs on respective chemokine or G-protein coupled receptor mediated functions, e.g., second messenger levels, i.e., Ca^{++}; inositol phosphate pool changes (see, e.g., Berridge (1993) Nature 361:315-325 or Billah and Anthes (1990) Biochem. J. 269:281-291); cellular morphology modification responses; phosphoinositide lipid turnover; an antiviral response. and others. Some detection methods allow for elimination of a separation step, e.g., a proximity sensitive detection system. Calcium sensitive dyes will be useful for detecting Ca^{++} levels, with a fluorimeter or a fluorescence cell sorting apparatus.

Rational drug design may also be based upon structural studies of the molecular shapes of the chemokines, other effectors or analogs, or the receptors. Effectors may be other proteins which

mediate other functions in response to ligand binding, or other proteins which normally interact with the receptor. One means for determining which sites interact with specific other proteins is a physical structure determination, e.g., x-ray crystallography or 2 dimensional NMR techniques. These will provide guidance as to which amino acid residues form molecular contact regions. For a detailed description of protein structural determination, see, e.g., Blundell and Johnson (1976) Protein Crystallography, Academic Press, New York.

Purified chemokine or receptor can be coated directly onto plates for use in the aforementioned drug screening techniques, and may be associated with detergents or lipids. However, non-neutralizing antibodies, e.g., to the chemokines or receptors can be used as capture antibodies to immobilize the respective protein on the solid phase.

Similar concepts also apply to the chemokine receptor embodiments of the invention.

IX. Kits

This invention also contemplates use of chemokine or receptor proteins, fragments thereof, peptides, binding compositions, and their fusion products in a variety of diagnostic kits and methods for detecting the presence of ligand, antibodies, or receptors. Typically the kit will have a compartment containing a defined chemokine or receptor peptide or gene segment or a reagent which recognizes one or the other, e.g., binding reagents.

A kit for determining the binding affinity of a test compound to a chemokine or receptor would typically comprise a test compound; a labeled compound, for example an antibody having known binding affinity for the protein; a source of chemokine or receptor (naturally occurring or recombinant); and a means for separating bound from free labeled compound, such as a solid phase for immobilizing the ligand or receptor. Once compounds are screened, those having suitable binding affinity to the ligand or receptor can be evaluated in suitable biological assays, as are well known in the art, to determine whether they act as agonists or

antagonists to the receptor. The availability of recombinant chemokine or receptor polypeptides also provide well defined standards for calibrating such assays or as positive control samples.

A preferred kit for determining the concentration of, for example, a chemokine or receptor in a sample would typically comprise a labeled compound, e.g., antibody, having known binding affinity for the target, a source of ligand or receptor (naturally occurring or recombinant) and a means for separating the bound from free labeled compound, for example, a solid phase for immobilizing the chemokine or receptor. Compartments containing reagents, and instructions for use or disposal, will normally be provided.

Antibodies, including antigen binding fragments, specific for the chemokine or receptor, or fragments are useful in diagnostic applications to detect the presence of elevated levels of chemokine, receptor, and/or its fragments. Such diagnostic assays can employ lysates, live cells, fixed cells, immunofluorescence, cell cultures, body fluids, and further can involve the detection of antigens related to the ligand or receptor in serum, or the like. Diagnostic assays may be homogeneous (without a separation step between free reagent and antigen complex) or heterogeneous (with a separation step). Various commercial assays exist, such as radioimmunoassay (RIA), enzyme-linked immunosorbent assay (ELISA), enzyme immunoassay (EIA), enzyme-multiplied immunoassay technique (EMIT), substrate-labeled fluorescent immunoassay (SLFIA), and the like. For example, unlabeled antibodies can be employed by using a second antibody which is labeled and which recognizes the primary antibody to a chemokine or receptor or to a particular fragment thereof. Similar assays have also been extensively discussed in the literature. See, e.g., Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH.

Anti-idiotypic antibodies may have similar uses to diagnose presence of antibodies against a chemokine or receptor, as such may be diagnostic of various abnormal states. For example, overproduction of a chemokine or receptor may result in production of various immunological reactions which may be diagnostic of

abnormal physiological states, particularly in various inflammatory or asthma conditions.

Frequently, the reagents for diagnostic assays are supplied in kits, so as to optimize the sensitivity of the assay. For the subject invention, depending upon the nature of the assay, the protocol, and the label, either labeled or unlabeled antibody or labeled chemokine or receptor is provided. This is usually in conjunction with other additives, such as buffers, stabilizers, materials necessary for signal production such as substrates for enzymes, and the like. Preferably, the kit will also contain instructions for proper use and disposal of the contents after use. Typically the kit has compartments or containers for each useful reagent. Desirably, the reagents are provided as a dry lyophilized powder, where the reagents may be reconstituted in an aqueous medium providing appropriate concentrations of reagents for performing the assay.

The aforementioned constituents of the drug screening and the diagnostic assays may be used without modification or may be modified in a variety of ways. For example, labeling may be achieved by covalently or non-covalently joining a moiety which directly or indirectly provides a detectable signal. In any of these assays, the ligand, test compound, chemokine, receptor, or antibodies thereto can be labeled either directly or indirectly. Possibilities for direct labeling include label groups: radiolabels such as ^{125}I, enzymes (U.S. Pat. No. 3,645,090) such as peroxidase and alkaline phosphatase, and fluorescent labels (U.S. Pat. No. 3,940,475) capable of monitoring the change in fluorescence intensity, wavelength shift, or fluorescence polarization. Possibilities for indirect labeling include biotinylation of one constituent followed by binding to avidin coupled to one of the above label groups.

There are also numerous methods of separating bound from the free ligand, or alternatively bound from free test compound. The chemokine or receptor can be immobilized on various matrixes, perhaps with detergents or associated lipids, followed by washing. Suitable matrixes include plastic such as an ELISA plate, filters, and beads. Methods of immobilizing the chemokine or receptor to a matrix include, without limitation, direct adhesion to plastic, use of

a capture antibody, chemical coupling, and biotin-avidin. The last step in this approach may involve the precipitation of antigen/antibody complex by any of several methods including those utilizing, e.g., an organic solvent such as polyethylene glycol or a salt such as ammonium sulfate. Other suitable separation techniques include, without limitation, the fluorescein antibody magnetizable particle method described in Rattle, et al. (1984) Clin. Chem. 30:1457-1461, and the double antibody magnetic particle separation as described in U.S. Pat. No. 4,659,678.

Methods for linking proteins or their fragments to the various labels have been extensively reported in the literature and do not require detailed discussion here. Many of the techniques involve the use of activated carboxyl groups either through the use of carbodiimide or active esters to form peptide bonds, the formation of thioethers by reaction of a mercapto group with an activated halogen such as chloroacetyl, or an activated olefin such as maleimide, for linkage, or the like. Fusion proteins will also find use in these applications.

Another diagnostic aspect of this invention involves use of oligonucleotide or polynucleotide sequences taken from the sequence of the chemokine or receptor. These sequences can be used as probes for detecting levels of the ligand message in samples from patients suspected of having an abnormal condition, e.g., an inflammatory, physiological, or developmental problem. The preparation of both RNA and DNA nucleotide sequences, the labeling of the sequences, and the preferred size of the sequences has received ample description and discussion in the literature. Normally an oligonucleotide probe should have at least about 14 nucleotides, usually at least about 18 nucleotides, and the polynucleotide probes may be up to several kilobases. Various labels may be employed, most commonly radionuclides, particularly ^{32}P. However, other techniques may also be employed, such as using biotin modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescers, enzymes, or the like. Alternatively,

antibodies may be employed which can recognize specific duplexes, including DNA duplexes, RNA duplexes, DNA-RNA hybrid duplexes, or DNA-protein duplexes. The antibodies in turn may be labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected. The use of probes to the novel anti-sense RNA may be carried out in conventional techniques such as nucleic acid hybridization, plus and minus screening, recombinational probing, hybrid released translation (HRT), and hybrid arrested translation (HART). This also includes amplification techniques such as polymerase chain reaction (PCR).

Diagnostic kits which also test for the qualitative or quantitative presence of other markers are also contemplated. Diagnosis or prognosis may depend on the combination of multiple indications used as markers. Thus, kits may test for combinations of markers. See, e.g., Viallet, et al. (1989) Progress in Growth Factor Res. 1:89-97.

X. Receptor for Chemokine; Ligands for Receptors

Having isolated a ligand binding partner of a specific interaction, methods exist for isolating the counter-partner. See, Gearing, et al EMBO J. 8:3667-4676 or McMahan, et al. (1991) EMBO J. 10:2821-2832. For example, means to label a chemokine without interfering with the binding to its receptor can be determined. For example, an affinity label can be fused to either the amino- or carboxy-terminus of the ligand. An expression library can be screened for specific binding of chemokine, e.g., by cell sorting, or other screening to detect subpopulations which express such a binding component. See, e.g., Ho, et al. (1993) Proc. Nat'l Acad. Sci. 90:11267-11271. Alternatively, a panning method may be used. See, e.g., Seed and Aruffo (1987) Proc. Nat'l. Acad. Sci. 84:3365-3369.

With a receptor, means to identify the ligand exist. Methods for using the receptor, e.g., on the cell membrane, can be used to screen for ligand by, e.g., assaying for a common G-protein linked signal such as Ca++ flux. See Lerner (1994) Trends in Neurosciences

17:142-146. It is likely that the ligands for these receptors are chemokines.

Protein cross-linking techniques with label can be applied to a isolate binding partners of a chemokine. This would allow identification of protein which specifically interacts with a chemokine, e.g., in a ligand-receptor like manner.

The broad scope of this invention is best understood with reference to the following examples, which are not intended to limit the invention to specific embodiments.

EXAMPLES

I. General Methods

Some of the standard methods are described or referenced, e.g., in Maniatis, et al. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual, (2d ed.), vols 1-3, CSH Press, NY; Ausubel, et al., Biology, Greene Publishing Associates, Brooklyn, NY; or Ausubel, et al. (1987 and Supplements) Current Protocols in Molecular Biology, Greene/Wiley, New York; Innis, et al. (eds.)(1990) PCR Protocols: A Guide to Methods and Applications Academic Press, N.Y. Methods for protein purification include such methods as ammonium sulfate precipitation, column chromatography, electrophoresis, centrifugation, crystallization, and others. See, e.g., Ausubel, et al. (1987 and periodic supplements); Deutscher (1990) "Guide to Protein Purification" in Methods in Enzymology, vol. 182, and other volumes in this series; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, Piscataway, N.J., or Bio-Rad, Richmond, CA. Combination with recombinant techniques allow fusion to appropriate segments, e.g., to a FLAG sequence or an equivalent which can be fused via a protease-removable sequence. See, e.g., Hochuli (1989) Chemische Industrie 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) Genetic Engineering, Principle and Methods 12:87-98, Plenum Press, N.Y.; and Crowe, et

al. (1992) QIAexpress: The High Level Expression & Protein Purification System QUIAGEN, Inc., Chatsworth, CA.

FACS analyses are described in Melamed, et al. (1990) Flow Cytometry and Sorting Wiley-Liss, Inc., New York, NY; Shapiro (1988) Practical Flow Cytometry Liss, New York, NY; and Robinson, et al. (1993) Handbook of Flow Cytometry Methods Wiley-Liss, New York, NY.

II. Isolation and characterization of chemokine cDNAs

A. Primate IBICK

The IBICK was isolated from a cDNA library made from a human astrocytoma cell line. See, e.g., Rani, et al. (1996) J. Biol. Chem. 271:22878-22884. There is reported a gene which is not inducible by IFN-α, but inducible by IFN-γ and/or IFN-β. Applicants have identified this gene as a chemokine, and designated it Interferon Beta Induced ChemoKine (IBICK), which is described in SEQ ID NO: 2. Individual cDNA clones are sequenced using standard methods, e.g., the Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA), and the sequence is further characterized.

The predicted signal sequence corresponds to amino acids met1 to about gly21, so the mature form should begin with phe22 and run about 74 amino acids. Additional processing may occur in a physiological system.

Computer analysis and alignments for related genes indicates the closest match is to the two IFN-γ regulated chemokines MIG and IP10, but other related molecules are chemoknes. See, e.g., Faubert (1993) Biochem. Biophys. Res. Commun. 192:223-230; and Luster, et al. (1985) Nature 315:672-676. This similarity in sequence may well correlate with similarity in regulation, which suggests related functions. The rarity of related sequences in the existing sequence databases suggests low message levels, tight negative regulation, and/or a distribution pattern in cell types not yet analysed. The IFN-γ regulatable nature of this chemokine suggests a role as an antiviral or antitumor agent. Its non-ELR chemokine structure suggests

angiostatic, in contrast to angiogenic, activity, which may be important in tumor therapy.

Other primate counterparts should be isolatable using the entire coding portion of this human clone as a hybridization probe. A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

B. Primate ILINCK

The ILINCK (SEQ ID NO: 3 and 4) was isolated from a cDNA library made from a human liver cell library. Total RNA can be isolated, e.g., using the guanidine thiocyanate/CsCl gradient procedure as described by Chirgwin, et al. (1978) Biochem. 18:5294-5299. Poly(A)+ RNA is isolated using, e.g., the OLIGOTEX mRNA isolation kit (QIAGEN). Such RNA from these cells is used to synthesize first strand cDNA, e.g., by using NotI/Oligo-dT primer (Gibco-BRL, Gaithersburg, MD). Double-stranded cDNA is synthesized, ligated with BstXI adaptors, digested with NotI, size fractionated for > 0.5 kilobase pairs (kb) and ligated into the NotI/BstXI sites of pJFE-14, a derivative of the pCDSRa vector. See Takebe, et al. (1985) Mol. Cell Biol. 8:466-472. Electro-competent E. coli DH10a cells (Gibco-BRL) are used for transformation.

The gene apparently produces at least two different sized transcripts, 0.6 kB and a 1.5 kB, which are differently regulated. The larger transcript is inducible by IL-10, which is unusual for a chemokine, so it has been designated InterLeukin 10 INduced ChemoKine (ILINCK), which is described in SEQ ID NO: 4. Individual cDNA clones were sequenced using standard methods, e.g., the Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA), and the sequence was further characterized.

The predicted signal sequence corresponds to amino acids met1 to about ser23, so the mature form should begin with gln24 and run about 73 amino acids. Additional processing may occur in a physiological system.

Computer analysis and alignments for related genes indicates the closest match is to other chemokines.

Other primate counterparts should be isolatable using the entire coding portion of this human clone as a hybridization probe. A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

The ILINCK mRNA is induced in monocytes by IL-10, a most notable feature. This observation strongly suggests that ILINCK has anti-inflammatory properties. It is postulated that ILINCK will be a potential therapeutic in autoimmune or other inflammatory disorders. See, e.g., Samter, et al. (eds) Immunological Diseases vols. 1 and 2, Little, Brown and Co.

C. Rodent CXC-143

The CXC-143 (SEQ ID NO: 5) was isolated from a cDNA library made from a mouse placenta cDNA library. The partial sequence provided lacks an identifiable initiation codon and termination codon. This chemokine has been designated CXC-143, and is described in SEQ ID NO: 6, 8 and 10. Individual cDNA clones were sequenced using standard methods, e.g., the Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA), and the sequence was further characterized, but the sequence remains incomplete. Clearly the chemokine is a non-ELR class CXC chemokine.

Computer analysis and alignments for related genes indicates the closest match is to other chemokine, the MIG, IP10, and the IBICK, all of which are IFN-γ inducible. This sequence similarity suggests a similar transcriptional regulation, and similar uses to the IBICK described above.

Other primate counterparts should be isolatable using the entire coding portion of this human clone as a hybridization probe. A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

D. Rodent MCP243

The MCP243 (SEQ ID NO: 11) was isolated from a cDNA library made from a mouse cDNA library. The partial sequence provided lacks an identifiable initiation codon and various upstream chemokine motifs. This chemokine has been designated MCP23, and is described in SEQ ID NO: 12 and 14. Individual cDNA clones are sequenced using standard methods, e.g., the Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA), and the sequence further characterized, but the sequence remains incomplete.

Computer analysis and alignments for related genes indicates the closest match is to various monocyte chemoattractant proteins. Clearly the encoded protein is a chemokine, and will have many similar biological activities related to the other members of the MCPs.

Other rodent counterparts should be isolatable using the entire coding portion of this mouse clone as a hybridization probe. A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

III. Isolation and characterization of GPCR cDNAs

A. Primate R277

The primate R277 clone was derived from human fetal tissue cDNA library. The nucleotide and amino acid sequences are provided in SEQ ID NO: 15, 16, 17 and 18.

Computer analysis suggests that the closest related genes are various G-protein coupled receptors. These include the chemokine receptors, and protease, e.g., thrombin, receptors. Structural motifs suggest that the receptor may contain motifs characteristic of the chemokine receptor family, and of the protease receptor family. The transmembrane segments, based upon hydrophobicity plots and comparisons with other similar GPCRs, should be about as follows for SEQ ID NO: 16: TM3 to val17; TM4 from arg36 to leu57; TM5 from asn89 to arg111; TM6 from leu134 to leu161; and TM7 from met178 to val200.

See, e.g., Loetscher, et al. (1996) J. Expt'l Med. 184:963-969. A DRY motif is found, e.g., near residue 18. The amino terminal segment is probably an extracellular segment (E1), and the others would be E2 between TM2 and TM3; E3 between TM4 and TM5; and E4 between TM6 and TM7. The intracellular segments should then run I1 between TM1 and TM2; I2 between TM3 and TM4, I3 between TM5 and TM6, and I4 the carboxy terminus from the end of TM7. Additional processing may occur in a physiological system. Conserved residues among the GPCRs would include, e.g., arg33, cys67, and gly217, among others in the transmembrane segments. A computer analysis of GPCR sequences will indicate residues characteristic of the family members. Core transmembrane segments for the R277 receptor sequence are predicted, using SEQ ID NO: 18 numbering, about: TM1 leu16 to leu41; TM2 ile51 to ile71; TM3 leu93 to leu125; TM4 leu132 to leu150; TM5 leu183 to val206; TM6 ile224 to ala256; and TM7 ile274 to val293.

Other primate counterparts should be isolatable using the entire coding portion of this human clone as a hybridization probe. A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

B. Rodent HST01.1

The rodent HST01.1 clone was derived from a cDNA library made from mouse TcR ab+ CD4- CD8- T cells. See Zlotnik, et al. (1992) J. Immunol. 149:1211-1215. Individual cDNA clones are sequenced using standard methods, and the sequence identified and further characterized. The partial nucleotide sequence is provided in SEQ ID NO: 19, encoding a polypeptide fragment of about 74 amino acids (SEQ ID NO: 20). Complete rodent HSTO1.1 nucleotide and amino acid sequences are prtovided in SEQ ID NO: 21 and 22.

Computer analysis suggests that the closest related genes are various G-protein coupled receptors. Structural motifs suggest that the receptor may contain motifs characteristic of the chemokine receptor family, and of the protease receptor family. In SEQ ID NO: 22, the transmembrane segments, based upon hydrophobicity plots and comparisons with other similar GPCRs, should be about as

follows: TM1 from leu58 to leu78; TM2 from phe90 to val110; TM3 from ala126 to phe146; TM4 from leu169 to leu189; TM5 from phe223 to val243; TM6 from val256 to asp277; and TM7 from val301 to fly321. A DRY box motif runs from about asp147 to ala155. Chemokine receptors are generally considered useful targets for novel drug discovery, where the therapeutics would agonise or antagonise the binding of natural ligand(s) of the receptor. These receptor-ligand interactions may result in inflammation, cell recruitment, an/or cell activation processes. Some of these receptors are the portal of entry of infectious agents, e.g., viruses. Therefore, therapeutics directed against the chomokine receptor may find application in these diseases. In addition, the receptors may be important in determining fundamental structure or physiological responses.

Other rodent counterparts should be isolatable using the entire coding portion of this mouse clone as a hybridization probe. A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

Ligand-receptor analysis has indicated that this receptor, when transfected into a cell, makes that cell responsive to the presence of mouse chemokines IP-10 (IFN-γ-Inducible Protein-10), MIG (monokine induced by IFN-g), and 6Ckine. See, e.g., Luster, et al. (1985) Nature 315:672-676; Ohmori and Hamilton (1990) Biochem. Biophys. Res. Commun. 168:1261-1267; Vanguri and Farber (1990) J. Biol. Chem. 265:15049-15057; Farber (1990) Proc. Nat'l Acad. Sci. USA 87:5238-5242; Farber (1993) Biochem. Biophys. Res. Commun. 192:223-230; and GenBank accession numbers AF006637; U88320, and U88322. However, the 6Ckine seems to bind differently from the MIG and IP-10, as it is incapable of desensitizing the response of the receptor to the other chemokines. MIG can desensitize the response to 6Ckine, but IP-10 does not. These results imply that 6Ckine may have angiostatic and antitumor activities similar to those of MIG and IP-10. See, e.g., Sgadari, et al. (1997) Blood 89:2635-2643; Arenberg, et al. (1996) J. Exp. Med. 184:981-992; Loetscher, et al. (1996)

J. Exp. Med. 184:963-969; Sgadari, et al. (1996) Proc. Natl Acad. Sci. USA 93:13791-13796; Angiolillo, et al. (1996) Ann. NY Acad. Sci. 795:158-167; Sarris, et al. (1996) Leukemia 10:757-765; Angiolillo, et al. (1995) J. Exp. Med. 182:155-162; Strieter, et al. (1995) J. Leukoc. Biol. 57:752-762; Strieter, et al. (1995) Biochem. Biophys. Res. Commun. 210:51-57; Clark-Lewis, et al. (1994) J. Biol. Chem. 269:16075-16081.

C. Rodent 941D12

The rodent 941D12 clone was derived from a cDNA library made from mouse Th3 polarized cells. Production of 3W Th1 or Th2 cells is described in Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367. Briefly, Th1 or Th2 populations were derived from CD4+ T cells stimulated with antigen and antigen presenting cells in the presence of IL-12 or IL-4. Cells were stimulated once each week for 3 weeks, then harvested and restimulated, e.g., with PMA and ionmycin for 4 h. See Murphy, et al. (1996) J. Exp. Med. 183:901-913. A subtraction step was introduced to remove sequences found in mouse L cells. See, e.g., Hara, et al. (1994) Blood 84:189-199.

Individual cDNA clones are sequenced using standard methods, and the sequence identified and further characterized. The partial nucleotide sequence is provided in SEQ ID NO: 23, encoding a polypeptide fragment of about 193 amino acids (SEQ ID NO: 24). Complete rodent 94ID12 nucleotide and amino acid sequences are provided in SEQ ID NO: 25 and 26.

Computer analysis suggests that the closest related genes are various G-protein coupled receptors. Structural motifs suggest that the receptor may contain motifs characteristic of the chemokine receptor family, and of the protease receptor family. The transmembrane segments, based upon hydrophobicity plots and comparisons with other similar GPCRs, should be about as follows on SEQ ID NO: 24: TM1 from val62 to phe88; TM2 from val98 to ala120; TM3 from val145 to leu157. For SEQ ID NO: 26, the predicted core transmembrane segments are about TM1 leu48 to ile64; TM2 val77 to leu93; TM3 val124 to val140; TM4 ile155 to val171; TM5 leu201 to ile217; TM6 ser238 to thr254; and TM7 ile279 to leu295.

Other rodent counterparts should be isolatable using the entire coding portion of this mouse clone as a hybridization probe.

A Southern blot may indicate the extent of homology across species, and either a cDNA library or mRNA can be screened to identify an appropriate cell source. The physiological state of many different cell types may also be evaluated for increased expression of the gene.

IV. Preparation of antibodies

Many standard methods are available for preparation of antibodies. For example, synthetic peptides may be prepared to be used as antigen, administered to an appropriate animal, and either polyclonal or monoclonal antibodies prepared. Short peptides, e.g., less than about 10 amino acids may be expressed as repeated sequences, while longer peptides may be used alone or conjugated to a carrier. For example, with the GPCRs, animals are immunized with peptides or complete proteins from SEQ ID NO: 12, 14, 16, or 18. Highest specificity will result when the polypeptides are selected from portions which are most unique, e.g., not from conserved sequence regions. The animals may be used to collect antiserum, or may be used to generate monoclonal antibodies.

Antiserum is evaluated for use, e.g., in an ELISA, and will be evaluated for utility in immunoprecipitation, e.g., typically native, or Western blot, e.g., denatured antigen, analysis. Monoclonal antibodies will also be evaluated for those same uses.

The antibodies provided will be useful as immunoaffinity reagents, as detection reagents, for immunohistochemistry, and as potential therapeutic reagents, either as agonist or antagonist reagents.

V. Assays for chemotactic activity of chemokines

Chemokine proteins are produced, e.g., in COS cells transfected with a plasmid carrying the chemokine cDNA by electroporation. See, Hara, et al. (1992) EMBO J. 10:1875-1884. Physical analytical methods may be applied, e.g., CD analysis, to compare tertiary structure to other chemokines to evaluate whether the protein has likely folded into an active conformation. After transfection, a culture supernatant is collected and subjected to bioassays. A mock control, e.g., a plasmid carrying the luciferase

cDNA, is used. See, de Wet, et al. (1987) Mol. Cell. Biol. 7:725-757. A positive control, e.g., recombinant murine MIP-1a from R&D Systems (Minneapolis, MN), is typically used. Likewise, antibodies may be used to block the biological activities, e.g., as a control.

Lymphocyte migration assays are performed as previously described, e.g., in Bacon, et al. (1988) Br. J. Pharmacol. 95:966-974. Murine Th2 T cell clones, CDC-25 (see Tony, et al. (1985) J. Exp. Med. 161:223-241) and HDK-1 (see Cherwinski, et al. (1987) J. Exp. Med. 166:1229-1244), made available from R. Coffman and A. O'Garra (DNAX, Palo Alto, CA), respectively, are used as controls.

Ca2+ flux upon chemokine stimulation is measured, e.g., according to the published procedure described in Bacon, et al. (1995) J. Immunol. 154:3654-3666.

Maximal numbers of migrating cells in response to the IBICK are measured. See Schall (1993) J. Exp. Med. 177:1821-1826. A dose-response curve is determined, preferably giving a characteristic bell shaped dose-response curve.

After stimulation with various chemokines, lymphocytes often exhibit a measurable intracellular Ca2+ flux. MIP-1a, e.g., is capable of inducing immediate transients of calcium mobilization. Typically, the levels of chemokine used in these assays will be comparable to those used for the chemotaxis assays (1/1000 dilution of conditioned supernatants).

Retroviral infection assays have also been described, and recent description of certain chemokine receptors in retroviral infection processes may indicate that similar roles may apply these receptors. See, e.g., Balter (1996) Science 272:1740 (describing evidence for chemokine receptors as coreceptors for HIV); and Deng, et al. (1996) Nature 381:661-666.

For receptors, biological activity may be measured in response to an appropriate ligand. The receptors are transfected into an assortment of cell types, each of which is likely to possess the intracellular signaling components compatible with the expressed receptor. Various ligand sources are tested to find a source of ligand which results in a G-protein coupled response. Alternatively, the cells are tested for Ca++ flux in response to such ligands. Flux may

be conveniently measured by electrophysiological means, or by Ca++ sensitive dyes.

VI. Analysis of individual variation

From the distribution data, an abundant easily accessible cell type is selected for sampling from individuals. Using PCR techniques, a large population of individuals are analysed for this gene. cDNA or other PCR methods are used to sequence the corresponding gene in the different individuals, and their sequences are compared. This indicates both the extent of divergence among racial or other populations, as well as determining which residues are likely to be modifiable without dramatic effects on function.

VII. Biological activities, direct and indirect

A robust and sensitive assay is selected as described above, e.g., on immune cells, neuronal cells, or stem cells. Chemokine is added to the assay in increasing doses to see if a dose response is detected. For example, in a proliferation assay, cells are plated out in plates. Appropriate culture medium is provided, and chemokine is added to the cells in varying amounts. Growth is monitored over a period of time which will detect either a direct effect on the cells, or an indirect effect of the chemokine.

Alternatively, an activation assay or attraction assay is used. An appropriate cell type is selected, e.g, hematopoietic cells, myeloid (macrophages, neutrophils, polymorphonuclear cells, etc.) or lymphoid (T cell, B cell, or NK cells), neural cells (neurons, neuroglia, oligodendrocytes, astrocytes, etc.), or stem cells, e.g., progenitor cells which differentiate to other cell types, e.g., gut crypt cells and undifferentiated cell types.

Other assays will be those which have been demonstrated with other chemokines. See, e.g., Schall and Bacon (1994) Current Opinion in Immunology 6:865-873; and Bacon and Schall (1996) Int. Arch. Allergy & Immunol. 109:97-109.

VIII. Structure activity relationship

Information on the criticality of particular residues is determined using standard procedures and analysis. Standard mutagenesis analysis is performed, e.g., by generating many different variants at determined positions, e.g., at the structural positions identified above, and evaluating biological activities of the variants. This may be performed to the extent of determining positions which modify activity, or to focus on specific positions to determine the residues which can be substituted to either retain, block, or modulate biological activity.

Alternatively, analysis of natural variants can indicate what positions tolerate natural mutations. This may result from populational analysis of variation among individuals, or across strains or species. Samples from selected individuals are analyzed, e.g., by PCR analysis and sequencing. This allows evaluation of population polymorphisms.

IX. Chromosomal localization

The cDNA is labeled, e.g., nick-translated with biotin-14 dATP and hybridized in situ at a final concentration of 5 ng/μl to metaphases from two normal males. Fluorescence in situ hybridization (FISH) method may be modified from that described by Callen, et al. (1990). Ann. Genet. 33:219-221, in that chromosomes are stained before analysis with both prodidium iodide (as counter stain) and DAPI (for chromosome identification). Images of metaphase preparations are captured by a CCD camera and computer enhanced. Identification of the approapriate labeled chromosomes is determined.

X. Expression analysis of chemokine/receptor genes

RNA blot and hybridization are performed according to the standard methods in Maniatis, et al. (1982) Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. An appropriate fragment or the whole coding sequence of a cDNA fragment is selected for use as a probe. To verify the amount of RNA loaded in each lane, the substrate membrane is

reprobed with a control cDNA, e.g., glyceraldehyde 3-phosphate dehydrogenase (G3PDH) cDNA (Clontech, Palo Alto CA).

Analysis of mRNA from the appropriate cell source using the probe will determine the natural size of message. It will also indicate whether different sized messages exist. The messages will be subject to analysis after isolation, e.g., by PCR or hybridization techniques.

Northern blot analysis may be performed on many different mRNA sources, e.g., different tissues, different species, or cells exhibiting defined physiological responses, e.g., activation conditions or developmental conditions. However, in certain cases, cDNA libraries may be used to evaluate sources which are difficult to prepare. A "reverse Northern" uses cDNA inserts removed from vector, but multiplicity of bands may reflect either different sized messages, or may be artifact due to incomplete reverse transcription in the preparation of the cDNA library. In such instances, verification may be appropriate by standard Northern analysis.

Similarly, Southern blots may be used to evaluate species distribution of a gene. The stringency of washes of the blot will also provide information as to the extent of homology of various species counterparts.

Tissue distribution, and cell distribution, may be evaluated by immunohistochemistry using antibodies. Alternatively, in situ nucleic acid hybridization may also be used in such analysis. Certain distribution data may be ascertained by the frequency and tissue types where messages have been found and collected in sequence databases, e.g., GenBank or proprietary collections.

A. IBICK

The IBICK was isolated from a human astrocyte cell. There is little distribution data generated at this time.

B. ILINCK

The IBICK was isolated from a human liver library. It is expressed in NK cells, gd T cells, and activated and resting monocytes. Libraries from human T cell lines, e.g., Mot81, HY106, and Mut72 show expression. Northern blots of adult spleen, thymus, prostate, testis, yterus, small intestine, colon, and peripheral

blood leukocytes give no detectable signal. These data suggest that the expression is in a limited subset of specialized cells. Neither activated splenocytes not activated PBLs give detectable signals.

C. CXC-143

The CXC-143 gene was identified from a cDNA library made from mouse placenta. The sequence has appeared in cDNA libraries from varous C57BL/6J mouse embryo or placenta. Notably, early embryo libraries, e.g., from 8.5 days post conception, exhibited signal. A homologous, human gene is found in placenta, fetal heart, breast, and fetal liver/spleen. This distibution suggests a role of the molecule in early stages of development. The molecule, or its antagonist, should be useful in various early developmental conditions.

D. R277 GPCR

The R277 gene was identified from a cDNA library made from 20 week human fetal liver/spleen tissues. Southern analysis in cDNA Libraries showed that the gene is highly expressed CD34 derived dendritic cells; and is also expressed in T & B cell libraries (varying degrees); activated monocytes; and NKL clone. On a Clontech Multiple Human Tissue blot, a transcript of about 2.4 kb was detected in spleen and PBL.

E. HST01.1 GPCR

The HST01.1 gene was isolated from an ab TCR+ CD4+ CD8+ cell library. Distribution analysis showed a strong positive signal by Northern analysis in spleen and lung. Southern analysis showed strong positive signals in Th1 clones, CD4+NK1.1+ cells, abTcR double negative cells, D1.1 resting Th1 T cells, and mesentary lymph nodes. Positive signals were detected in macrophages, Th2 cells, and CD3+abTcR double negative cells, D1.1 ConA stimulated Th1 T cells, resting J774 cells, and LPS and IL-10 stimulated J774 cells. Weaker signals were detected in thymus, activated pro-T cells, and LPS and anti-IL-10 treated J774 cells. Very weak signals were detected in large B cell line from spleen, from dendritic cells from a resting spleen, and colon. No signal was detected in resting pro-T cells, CD44-CD25+ pre-T cells, CD35+ resting Th2 T cells, ConA stimulated CD35+ Th2 T cells, mature B cell leukemia, CH12 B cell line, B cells

from LPS treated spleen, resting dendritic cells from bone marrow, RAW264.7 moncyte cell line, and total Peyer's patch.

F. 941D12 GPCR

The 941D12 GPCR was isolated from a mouse 3 week polarized Th2 cell cDNA library, subtracted with cDNA sequences from a mouse L cell. See, e.g., Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; and Murphy, et al. (1996) J. Exp. Med. 183:901-913.

Southern Analysis: DNA (5 mg) from a primary amplified cDNA library was digested with appropriate restriction enzymes to release the inserts, run on a 1% agarose gel and transferred to a nylon membrane (Schleicher and Schuell, Keene, NH).

Samples for mouse mRNA isolation include, e.g.: resting mouse fibroblastic L cell line (C200); Braf:ER (Braf fusion to estrogen receptor) transfected cells, control (C201); T cells, TH1 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IFN-g and anti IL-4; T200); T cells, TH2 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IL-4 and anti-IFN-γ; T201); T cells, highly TH1 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T202); T cells, highly TH2 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T203); CD44- CD25+ pre T cells, sorted from thymus (T204); TH1 T cell clone D1.1, resting for 3 weeks after last stimulation with antigen (T205); TH1 T cell clone D1.1, 10 mg/ml ConA stimulated 15 h (T206); TH2 T cell clone CDC35, resting for 3 weeks after last stimulation with antigen (T207); TH2 T cell clone CDC35, 10 mg/ml ConA stimulated 15 h (T208); Mel 14+ naive T cells from spleen, resting (T209); Mel14+ T cells, polarized to Th1 with IΦN–γ/IL-12/anti-IL-4 for 6, 12, 24 h pooled (T210); Mel 14+ T cells, polarized to Th2 with IL-4/anti-IΦN–γ for 6, 13, 24 h pooled (T211); unstimulated mature B cell leukemia cell line A20 (B200); unstimulated B cell line CH12 (B201); unstimulated large B cells from spleen (B202); B cells from total spleen, LPS activated (B203); metrizamide enriched dendritic cells from spleen, resting (D200); dendritic cells from bone marrow, resting (D201); monocyte cell line RAW 264.7 activated with LPS 4 h (M200); bone-marrow macrophages derived with GM

and M-CSF (M201); macrophage cell line J774, resting (M202); macrophage cell line J774 + LPS + anti-IL-10 at 0.5, 1, 3, 6, 12 h pooled (M203); macrophage cell line J774 + LPS + IL-10 at 0.5, 1, 3, 5, 12 h pooled(M204); aerosol challenged mouse lung tissue, Th2 primers, aerosol OVA challenge 7, 14, 23 h pooled (see Garlisi, et al. (1995) Clinical Immunology and Immunopathology 75:75-83; X206); Nippostrongulus-infected lung tissue (see Coffman, et al. (1989) Science 245:308-310; X200); total adult lung, normal (O200); total lung, rag-1 (see Schwarz, et al. (1993) Immunodeficiency 4:249-252; O205); IL-10 K.O. spleen (see Kuhn, et al. (1991) Cell 75:263-274; X201); total adult spleen, normal (O201); total spleen, rag-1 (O207); IL-10 K.O. Peyer's patches (O202); total Peyer's patches, normal (O210); IL-10 K.O. mesenteric lymph nodes (X203); total mesenteric lymph nodes, normal (O211); IL-10 K.O. colon (X203); total colon, normal (O212); NOD mouse pancreas (see Makino, et al. (1980) Jikken Dobutsu 29:1-13; X205); total thymus, rag-1 (O208); total kidney, rag-1 (O209); total heart, rag-1 (O202); total brain, rag-1 (O203); total testes, rag-1 (O204); total liver, rag-1 (O206); rat normal joint tissue (O300); and rat arthritic joint tissue (X300).

High signals were detected in IL-10 K.O. Peyer's patches (O202); total Peyer's patches, normal (O210); TH2 T cell clone CDC35, resting for 3 weeks after last stimulation with antigen (T207); T cells, highly TH2 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T203); total kidney, rag-1 (O209); and total heart, rag-1 (O202). Significant signals were detected in bone-marrow macrophages derived with GM and M-CSF (M201); T cells, TH2 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IL-4 and anti-IΦN–γ; T201); dendritic cells from bone marrow, resting (D201); total brain, rag-1 (O203); total liver, rag-1 (O206); total colon, normal (O212); and total thymus, rag-1 (O208). Weak signals were detected in TH2 T cell clone CDC35, 10 mg/ml ConA stimulated 15 h (T208); macrophage cell line J774, resting (M202); IL-10 K.O. colon (X203); Mel 14+ T cells, polarized to Th2 with IL-4/anti-IFN-γ for 6, 13, 24 h pooled (T211); macrophage cell line J774 + LPS + anti-IL-10 at 0.5, 1, 3, 6, 12 h pooled (M203); TH1 T cell clone D1.1, 10 mg/ml ConA stimulated 15 h (T206); total lung,

rag-1 (see Schwarz, et al. (1993) Immunodeficiency 4:249-252; O205); total spleen, rag-1 (O207); IL-10 K.O. mesenteric lymph nodes (X203); total testes, rag-1 (O204); T cells, TH1 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IΦN-γ and anti IL-4; T200); TH1 T cell clone D1.1, resting for 3 weeks after last stimulation with antigen (T205); monocyte cell line RAW 264.7 activated with LPS 4 h (M200); and Nippostrongulus-infected lung tissue (see Coffman, et al. (1989) Science 245:308-310; X200). Other samples in the list gave no detectable signal.

XI. Screening for receptor/ligand

Labeled reagent is useful for screening of an expression library made from a cell line which expresses a chemokine or receptor, as appropriate. Standard staining techniques are used to detect or sort intracellular or surface expressed ligand, or surface expressing transformed cells are screened by panning. Screening of intracellular expression is performed by various staining or immunofluorescence procedures. See also, e.g., McMahan, et al. (1991) EMBO J. 10:2821-2832.

For example, on day 0, precoat 2-chamber permanox slides with 1 ml per chamber of fibronectin, 10 ng/ml in PBS, for 30 min at room temperature. Rinse once with PBS. Then plate COS cells at 2-3 x 10^5 cells per chamber in 1.5 ml of growth media. Incubate overnight at 37° C.

On day 1 for each sample, prepare 0.5 ml of a solution of 66 mg/ml DEAE-dextran, 66 mM chloroquine, and 4 mg DNA in serum free DME. For each set, a positive control is prepared, e.g., of huIL-10-FLAG cDNA at 1 and 1/200 dilution, and a negative mock. Rinse cells with serum free DME. Add the DNA solution and incubate 5 hr at 37° C. Remove the medium and add 0.5 ml 10% DMSO in DME for 2.5 min. Remove and wash once with DME. Add 1.5 ml growth medium and incubate overnight.

On day 2, change the medium. On days 3 or 4, the cells are fixed and stained. Rinse the cells twice with Hank's Buffered Saline Solution (HBSS) and fix in 4% paraformaldehyde (PFA)/glucose for 5 min. Wash 3X with HBSS. The slides may be stored at -80° C after

all liquid is removed. For each chamber, 0.5 ml incubations are performed as follows. Add HBSS/saponin(0.1%) with 32 ml/ml of 1M NaN_3 for 20 min. Cells are then washed with HBSS/saponin 1X. Add antibody complex to cells and incubate for 30 min. Wash cells twice with HBSS/saponin. Add second antibody, e.g., Vector anti-mouse antibody, at 1/200 dilution, and incubate for 30 min. Prepare ELISA solution, e.g., Vector Elite ABC horseradish peroxidase solution, and preincubate for 30 min. Use, e.g., 1 drop of solution A (avidin) and 1 drop solution B (biotin) per 2.5 ml HBSS/saponin. Wash cells twice with HBSS/saponin. Add ABC HRP solution and incubate for 30 min. Wash cells twice with HBSS, second wash for 2 min, which closes cells. Then add Vector diaminobenzoic acid (DAB) for 5 to 10 min. Use 2 drops of buffer plus 4 drops DAB plus 2 drops of H_2O_2 per 5 ml of glass distilled water. Carefully remove chamber and rinse slide in water. Air dry for a few minutes, then add 1 drop of Crystal Mount and a cover slip. Bake for 5 min at 85-90° C.

Alternatively, the binding compositions are used to affinity purify or sort out cells expressing the ligand or receptor. See, e.g., Sambrook, et al. or Ausubel et al.

All references cited herein are incorporated herein by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Many modification an variations of this invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited only by the terms of the appended claims, along with the full scope of the equivalents to which such claims are entitled.

SEQUENCE LISTING

SEQ ID NO: 1 is primate IBICK nucleotide sequence.
SEQ ID NO: 2 is primate IBICK amino acid sequence.
SEQ ID NO: 3 is primate ILINCK nucleotide sequence.
SEQ ID NO: 4 is primate ILINCK amino acid sequence.
SEQ ID NO: 5 is partial rodent CXC-143 nucleotide sequence.
SEQ ID NO: 6 is partial rodent CXC-143 amino acid sequence.
SEQ ID NO: 7 is partial alternative rodent CXC-143 nucleotide sequence.
SEQ ID NO: 8 is partial alternative rodent CXC-143 amino acid sequence.
SEQ ID NO: 9 is revised partial rodent CXC-143 nucleotide sequence.
SEQ ID NO: 10 is revised partial rodent CXC-143 amino acid sequence.
SEQ ID NO: 11 is partial rodent MCP243 nucleotide sequence.
SEQ ID NO: 12 is partial rodent MCP243 amino acid sequence.
SEQ ID NO: 13 is revised complete rodent MCP243 nucleotide sequence.
SEQ ID NO: 14 is revised complete rodent MCP243 amino acid sequence.
SEQ ID NO: 15 is primate R277 nucleotide sequence.
SEQ ID NO: 16 is primate R277 amino acid sequence.
SEQ ID NO: 17 is revised primate R277 nucleotide sequence.
SEQ ID NO: 18 is revised primate R277 amino acid sequence.
SEQ ID NO: 19 is partial rodent HST01.1 nucleotide sequence.
SEQ ID NO: 20 is partial rodent HST01.1 amino acid sequence.
SEQ ID NO: 21 is rodent HST01.1 nucleotide sequence.
SEQ ID NO: 22 is rodent HST01.1 amino acid sequence.
SEQ ID NO: 23 is partial rodent 941D12 nucleotide sequence.
SEQ ID NO: 24 is partial rodent 941D12 amino acid sequence.
SEQ ID NO: 25 is rodent 941D12 nucleotide sequence.
SEQ ID NO: 26 is rodent 941D12 amino acid sequence.

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Schering Corp.
(B) STREET: 2000 Galloping Hill Road
(C) CITY: Kenilworth
(D) STATE: New Jersey
(F) ZIP: 07033-0530
(G) TELEPHONE: (908) 298-5056
(H) TELEFAX: (908) 298-5388

(ii) TITLE OF INVENTION: Mammalian Chemokines; Receptors; Reagents; Uses

(iii) NUMBER OF SEQUENCES: 26

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: Macintosh
(C) OPERATING SYSTEM: Macintosh OS 8.0
(D) SOFTWARE: Microsoft Word 5.1

(v) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 22-JAN-1998

(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/036,715
(B) FILING DATE: 23-JAN-1997

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 246
(D) OTHER INFORMATION: /note= "nucleotide 246 designated C, may be A, C, G, or T"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 23..307

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 86..307

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GCAGCAACAG CAAAAAACAA AC ATG AGT GTG AAG GGC ATG GCT ATA GCC TTG
52
                         Met Ser Val Lys Gly Met Ala Ile Ala Leu
                         -21 -20                 -15

GCT GTG ATA TTG TGT GCT ACA GTT GTT CAA GGC TTC CCC ATG TTC AAA
100
Ala Val Ile Leu Cys Ala Thr Val Val Gln Gly Phe Pro Met Phe Lys
    -10                 -5                  1               5

AGA GGA CGC TGT CTT TGC ATA GGC CCT GGG GTA AAA GCA GTG AAA GTG
148
Arg Gly Arg Cys Leu Cys Ile Gly Pro Gly Val Lys Ala Val Lys Val
                10                  15                  20

GCA GAT ATT GAG AAA GCC TCC ATA ATG TAC CCA AGT AAC AAC TGT GAC
196
Ala Asp Ile Glu Lys Ala Ser Ile Met Tyr Pro Ser Asn Asn Cys Asp
            25                  30                  35
```

AAA ATA GAA GTG ATT ATT ACC CTG AAA GAA ATA AGG AAA CGA TGC CTA
244
Lys Ile Glu Val Ile Ile Thr Leu Lys Glu Ile Arg Lys Arg Cys Leu
40 45 50

ACT CCA AAT CGA GCA TCG AAG CAA GCA AGG CTT ATA ATC AAA AAA GCT
292
Thr Pro Asn Arg Ala Ser Lys Gln Ala Arg Leu Ile Ile Lys Lys Ala
55 60 65

GAA AGA AAG AAT TTT TGAAAATATC AAAACATATG AAGTCCTGGA AAAGGGCATC
347
Glu Arg Lys Asn Phe
70

TGAAAAACCT AGAACAAGAT TAACTGTGAC TACTGAAATG ACAAGAATTC TACAGTAGGA
407

AACTGAGACT TTTCTATGGT TTTGTGACTT TCAACTTTTG TACAGTTATG TGAAGGATGA
467

A
468

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 95 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Val Lys Gly Met Ala Ile Ala Leu Ala Val Ile Leu Cys Ala
-21 -20 -15 -10

Thr Val Val Gln Gly Phe Pro Met Phe Lys Arg Gly Arg Cys Leu Cys
-5 1 5 10

Ile Gly Pro Gly Val Lys Ala Val Lys Val Ala Asp Ile Glu Lys Ala
15 20 25

Ser Ile Met Tyr Pro Ser Asn Asn Cys Asp Lys Ile Glu Val Ile Ile
30 35 40

Thr Leu Lys Glu Ile Arg Lys Arg Cys Leu Thr Pro Asn Arg Ala Ser
45 50 55

Lys Gln Ala Arg Leu Ile Ile Lys Lys Ala Glu Arg Lys Asn Phe

60 65 70

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 503 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 51..410

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 120..410

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GGGCTGCAGG  AATTCGGCAC  GAGCTGAAGC  TGTACTGCCT  CGCTGAGAGG  ATG  AAG
56
                                                            Met Lys
                                                            -23

GTC TCC GAG GCT GCC CTG TCT CTC CTT GTC CTC ATC CTT ATC ATT ACT
104
Val Ser Glu Ala Ala Leu Ser Leu Leu Val Leu Ile Leu Ile Ile Thr
    -20                 -15                 -10

TCG GCT TCT CGC AGC CAG CCA AAA GTT CCT GAG TGG GTG AAC ACC CCA
152
Ser Ala Ser Arg Ser Gln Pro Lys Val Pro Glu Trp Val Asn Thr Pro
 -5                  1               5                  10

TCC ACC TGC TGC CTG AAG TAT TAT GAG AAA GTG TTG CCA AGG AGA CTA
200
Ser Thr Cys Cys Leu Lys Tyr Tyr Glu Lys Val Leu Pro Arg Arg Leu
             15                  20                  25

GTG GTG GGA TAC AGA AAG GCC CTC AAC TGT CAC CTG CCA GCA ATC ATC
248
Val Val Gly Tyr Arg Lys Ala Leu Asn Cys His Leu Pro Ala Ile Ile
         30                  35                  40

TTC GTC ACC AAG AGG AAC CGA GAA GTC TGC ACC AAC CCC AAT GAC GAC
296
Phe Val Thr Lys Arg Asn Arg Glu Val Cys Thr Asn Pro Asn Asp Asp
     45                  50                  55
```

TGG GTC CAA GAG TAC ATC AAG GAT CCC AAC CTA CCT TTG CTG CCT ACC
344
Trp Val Gln Glu Tyr Ile Lys Asp Pro Asn Leu Pro Leu Leu Pro Thr
60 65 70 75

AGG AAC TTG TCC ACG GTT AAA ATT ATT ACA GCA AAG AAT GGT CAA CCC
392
Arg Asn Leu Ser Thr Val Lys Ile Ile Thr Ala Lys Asn Gly Gln Pro
80 85 90

CAG CTC CTC AAC TCC CAG TGATGACAAG CTTTAGTGGA AGCCCTTGTT
440
Gln Leu Leu Asn Ser Gln
95

TACAGAAAAA AAGGGTTAAC CTATGAAAAC AGGGGAAGCC TTATTTAGCT GAAACTAACC
500

CTC
503

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Lys Val Ser Glu Ala Ala Leu Ser Leu Leu Val Leu Ile Leu Ile
-23 -20 -15 -10

Ile Thr Ser Ala Ser Arg Ser Gln Pro Lys Val Pro Glu Trp Val Asn
-5 1 5

Thr Pro Ser Thr Cys Cys Leu Lys Tyr Tyr Glu Lys Val Leu Pro Arg
10 15 20 25

Arg Leu Val Val Gly Tyr Arg Lys Ala Leu Asn Cys His Leu Pro Ala
30 35 40

Ile Ile Phe Val Thr Lys Arg Asn Arg Glu Val Cys Thr Asn Pro Asn
45 50 55

Asp Asp Trp Val Gln Glu Tyr Ile Lys Asp Pro Asn Leu Pro Leu Leu
60 65 70

Pro Thr Arg Asn Leu Ser Thr Val Lys Ile Ile Thr Ala Lys Asn Gly

```
          75                  80                  85

     Gln Pro Gln Leu Leu Asn Ser Gln
      90                  95
5
     (2) INFORMATION FOR SEQ ID NO:5:

          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 375 base pairs
10             (B) TYPE: nucleic acid
               (C) STRANDEDNESS: single
               (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: cDNA
15

          (ix) FEATURE:
               (A) NAME/KEY: CDS
               (B) LOCATION: 3..311
20
          (ix) FEATURE:
               (A) NAME/KEY: -
               (B) LOCATION: 1

25        (ix) FEATURE:
               (A) NAME/KEY: unsure
               (B) LOCATION: 211
               (D) OTHER INFORMATION: /note= "nucleotide 211 may be
       absent; corresponding sequences presented as SEQ ID NO: 7 and 8"
30
          (ix) FEATURE:
               (A) NAME/KEY: -
               (B) LOCATION: 1

35
          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

     GC TCC TGC TGC TCC TCG GCT GTC GCC TCG CGC GTG GAC GGG TCC AAG
     47
40      Ser Cys Cys Ser Ser Ala Val Ala Ser Arg Val Asp Gly Ser Lys
         1               5                  10                  15

     TGT AAG TGT TCC CGG AAG GGG CCC AAG ATC CGC TAC AGC GAC GTG AAG
     95
45   Cys Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr Ser Asp Val Lys
                     20                  25                  30

     AAG CTG GAA ATG AAG CCA AAG TAC CCA CAC TGC GAG GAG AAG ATG GTT
     143
50   Lys Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu Glu Lys Met Val
                35                  40                  45
```

```
ATC GTC ACC ACC AAG AGC ATG TCC AGG TAC CGG GGC CAG GAG CAC TGC
191
Ile Val Thr Thr Lys Ser Met Ser Arg Tyr Arg Gly Gln Glu His Cys
         50                  55                  60

CTG CAC CCT AAG CTG CAG AGC ACC AAA CGC TTC ATC AAG TGG TAC AAT
239
Leu His Pro Lys Leu Gln Ser Thr Lys Arg Phe Ile Lys Trp Tyr Asn
     65                  70                  75

GCC TGG AAC GAG AAG CGC AGG GTC TAC GAA GAA TAG GGT GGA CGA TCA
287
Ala Trp Asn Glu Lys Arg Arg Val Tyr Glu Glu  *  Gly Gly Arg Ser
 80                  85                  90                  95

TGG AAA GAA AAA CTC CAG GCC AGT TGAGAGACTT CAGCAGAGGA CTTTGCAGAT
341
Trp Lys Glu Lys Leu Gln Ala Ser
                100

TAAAATAAAA            GCCCTTTCTT            TCTCACAAGC            ATAA
375
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Ser Cys Cys Ser Ser Ala Val Ala Ser Arg Val Asp Gly Ser Lys Cys
  1               5                  10                  15

Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr Ser Asp Val Lys Lys
             20                  25                  30

Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu Glu Lys Met Val Ile
         35                  40                  45

Val Thr Thr Lys Ser Met Ser Arg Tyr Arg Gly Gln Glu His Cys Leu
     50                  55                  60

His Pro Lys Leu Gln Ser Thr Lys Arg Phe Ile Lys Trp Tyr Asn Ala
 65                  70                  75                  80

Trp Asn Glu Lys Arg Arg Val Tyr Glu Glu  *  Gly Gly Arg Ser Trp
                 85                  90                  95
```

```
Lys Glu Lys Leu Gln Ala Ser
            100

(2) INFORMATION FOR SEQ ID NO:7:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 374 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA


    (ix) FEATURE:
          (A) NAME/KEY: CDS
          (B) LOCATION: 3..371

    (ix) FEATURE:
          (A) NAME/KEY: CDS
          (B) LOCATION: 3..371


    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GC TCC TGC TGC TCC TCG GCT GTC GCC TCG CGC GTG GAC GGG TCC AAG
47
   Ser Cys Cys Ser Ser Ala Val Ala Ser Arg Val Asp Gly Ser Lys
     1               5                  10                  15

TGT AAG TGT TCC CGG AAG GGG CCC AAG ATC CGC TAC AGC GAC GTG AAG
95
Cys Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr Ser Asp Val Lys
                20                  25                  30

AAG CTG GAA ATG AAG CCA AAG TAC CCA CAC TGC GAG GAG AAG ATG GTT
143
Lys Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu Glu Lys Met Val
            35                  40                  45

ATC GTC ACC ACC AAG AGC ATG TCC AGG TAC CGG GGC CAG GAG CAC TGC
191
Ile Val Thr Thr Lys Ser Met Ser Arg Tyr Arg Gly Gln Glu His Cys
        50                  55                  60

CTG CAC CCT AAG CTG CAG ACA CCA AAC GCT TCA TCA AGT GGT ACA ATG
239
Leu His Pro Lys Leu Gln Thr Pro Asn Ala Ser Ser Ser Gly Thr Met
    65                  70                  75

CCT GGA ACG AGA AGC GCA GGG TCT ACG AAG AAT AGG GTG GAC GAT CAT
287
Pro Gly Thr Arg Ser Ala Gly Ser Thr Lys Asn Arg Val Asp Asp His
```

```
      80                  85                  90                  95

    GGA AAG AAA AAC TCC AGG CCA GTT GAG AGA CTT CAG CAG AGG ACT TTG
    335
 5  Gly Lys Lys Asn Ser Arg Pro Val Glu Arg Leu Gln Gln Arg Thr Leu
                   100                 105                 110

    CAG  ATT  AAA  ATA  AAA  GCC  CTT  TCT  TTC  TCA  CAA  GCA  TAA
    374
10  Gln Ile Lys Ile Lys Ala Leu Ser Phe Ser Gln Ala
            115                 120


    (2) INFORMATION FOR SEQ ID NO:8:
15
          (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 123 amino acids
                (B) TYPE: amino acid
                (D) TOPOLOGY: linear
20
          (ii) MOLECULE TYPE: protein

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

25  Ser Cys Cys Ser Ser Ala Val Ala Ser Arg Val Asp Gly Ser Lys Cys
     1               5                  10                  15

    Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr Ser Asp Val Lys Lys
                20                  25                  30
30
    Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu Glu Lys Met Val Ile
            35                  40                  45

    Val Thr Thr Lys Ser Met Ser Arg Tyr Arg Gly Gln Glu His Cys Leu
35       50                  55                  60

    His Pro Lys Leu Gln Thr Pro Asn Ala Ser Ser Ser Gly Thr Met Pro
     65                  70                  75                  80

40  Gly Thr Arg Ser Ala Gly Ser Thr Lys Asn Arg Val Asp Asp His Gly
                    85                  90                  95

    Lys Lys Asn Ser Arg Pro Val Glu Arg Leu Gln Gln Arg Thr Leu Gln
                100                 105                 110
45
    Ile Lys Ile Lys Ala Leu Ser Phe Ser Gln Ala
            115                 120

    (2) INFORMATION FOR SEQ ID NO:9:
50
         (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 276 base pairs
               (B) TYPE: nucleic acid
```

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..273

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CTC CTG CTG CTC CTC GCG CTG TAC GCC TCG CGC GTG GAC GGG TCC AAG
48
Leu Leu Leu Leu Leu Ala Leu Tyr Ala Ser Arg Val Asp Gly Ser Lys
 1               5                  10                  15

TGT AAG TGT TCC CGG AAG GGG CCC AAG ATC CGC TAC AGC GAC GTG AAG
96
Cys Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr Ser Asp Val Lys
            20                  25                  30

AAG CTG GAA ATG AAG CCA AAG TAC CCA CAC TGC GAG GAG AAG ATG GTT
144
Lys Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu Glu Lys Met Val
        35                  40                  45

ATC GTC ACC ACC AAG AGC ATG TCC AGG TAC CGG GGC CAG GAG CAC TGC
192
Ile Val Thr Thr Lys Ser Met Ser Arg Tyr Arg Gly Gln Glu His Cys
    50                  55                  60

CTG CAC CCT AAG CTG CAG AGC ACC AAA CGC TTC ATC AAG TGG TAC AAT
240
Leu His Pro Lys Leu Gln Ser Thr Lys Arg Phe Ile Lys Trp Tyr Asn
65                  70                  75                  80

GCC  TGG  AAC  GAG  AAG  CGC  AGG  GTC  TAC  GAA  GAA  TAG
276
Ala Trp Asn Glu Lys Arg Arg Val Tyr Glu Glu
                85                  90
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Leu Leu Leu Leu Leu Ala Leu Tyr Ala Ser Arg Val Asp Gly Ser Lys
 1               5                  10                  15

Cys Lys Cys Ser Arg Lys Gly Pro Lys Ile Arg Tyr Ser Asp Val Lys
            20                  25                  30

Lys Leu Glu Met Lys Pro Lys Tyr Pro His Cys Glu Glu Lys Met Val
        35                  40                  45

Ile Val Thr Thr Lys Ser Met Ser Arg Tyr Arg Gly Gln Glu His Cys
    50                  55                  60

Leu His Pro Lys Leu Gln Ser Thr Lys Arg Phe Ile Lys Trp Tyr Asn
65                  70                  75                  80

Ala Trp Asn Glu Lys Arg Arg Val Tyr Glu Glu
                85                  90
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 315 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..162

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
AGA AGT TCG GAA CTG AGG GAG AGA ATC AAC AAT ATC CAG TGC CCC ATG
48
Arg Ser Ser Glu Leu Arg Glu Arg Ile Asn Asn Ile Gln Cys Pro Met
 1               5                  10                  15

GAA GCT GTG GTT TTC CAG ACC AAG CAG GGT ATG TCT CTC TGT GTA GAC
96
Glu Ala Val Val Phe Gln Thr Lys Gln Gly Met Ser Leu Cys Val Asp
            20                  25                  30

CCC ACA CAG AAG TGG GTC AGT GAG TAC ATG GAG ATC CTT GAC CAG AAG
144
Pro Thr Gln Lys Trp Val Ser Glu Tyr Met Glu Ile Leu Asp Gln Lys
        35                  40                  45
```

```
     TCT   CAA   ATT   CTG   CAG   CCT   TGAACCTTCA   CACCTGAGTT   AAGAGACAGC
     192
     Ser Gln Ile Leu Gln Pro
          50
5
     CAAAGCTGGA  AGTTCTCCCC  TAATCTTCTC  CAGGCAGAGA  GATGTTACAA  GCAGATGGTG
     252

     CCTGGGCTGC  GTGTTTTCTC  ATCCTTGTCT  GTTATATGAA  CAACTGAAAT  AAAAGCTTAC
10   312

     ACT
     315

15
     (2) INFORMATION FOR SEQ ID NO:12:

           (i) SEQUENCE CHARACTERISTICS:
                 (A) LENGTH: 54 amino acids
20               (B) TYPE: amino acid
                 (D) TOPOLOGY: linear

          (ii) MOLECULE TYPE: protein

25        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

     Arg Ser Ser Glu Leu Arg Glu Arg Ile Asn Asn Ile Gln Cys Pro Met
       1               5                  10                  15

30   Glu Ala Val Val Phe Gln Thr Lys Gln Gly Met Ser Leu Cys Val Asp
                  20                  25                  30

     Pro Thr Gln Lys Trp Val Ser Glu Tyr Met Glu Ile Leu Asp Gln Lys
              35                  40                  45
35
     Ser Gln Ile Leu Gln Pro
          50

     (2) INFORMATION FOR SEQ ID NO:13:
40
          (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 294 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
45              (D) TOPOLOGY: linear

         (ii) MOLECULE TYPE: cDNA


50       (ix) FEATURE:
                (A) NAME/KEY: CDS
                (B) LOCATION: 1..291
```

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 58..291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
ATG AAG ATC TAC GCA GTG CTT CTT TGC CTG CTG CTC ATA GCT GTC CCT
48
Met Lys Ile Tyr Ala Val Leu Leu Cys Leu Leu Leu Ile Ala Val Pro
-19             -15                 -10                 -5

GTC AGC CCA GAG AAG CTG ACT GGG CCA GAT AAG GCT CCA GTC ACC TGC
96
Val Ser Pro Glu Lys Leu Thr Gly Pro Asp Lys Ala Pro Val Thr Cys
                1               5                   10

TGC TTT CAT GTA CTA AAG CTG AAG ATC CCC CTT CGG GTG CTG AAA AGC
144
Cys Phe His Val Leu Lys Leu Lys Ile Pro Leu Arg Val Leu Lys Ser
     15                  20                  25

TAC GAG AGA ATC AAC AAT ATC CAG TGC CCC ATG GAA GCT GTG GTT TTC
192
Tyr Glu Arg Ile Asn Asn Ile Gln Cys Pro Met Glu Ala Val Val Phe
 30                  35                  40                  45

CAG ACC AAG CAG GGT ATG TCT CTC TGT GTA GAC CCC ACA CAG AAG TGG
240
Gln Thr Lys Gln Gly Met Ser Leu Cys Val Asp Pro Thr Gln Lys Trp
                50                  55                  60

GTC AGT GAG TAC ATG GAG ATC CTT GAC CAG AAG TCT CAA ATT CTG CAG
288
Val Ser Glu Tyr Met Glu Ile Leu Asp Gln Lys Ser Gln Ile Leu Gln
            65                  70                  75

CCT                                                         TGA
294
Pro
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Met Lys Ile Tyr Ala Val Leu Leu Cys Leu Leu Leu Ile Ala Val Pro
-19             -15                 -10                  -5

Val Ser Pro Glu Lys Leu Thr Gly Pro Asp Lys Ala Pro Val Thr Cys
               1                5                  10

Cys Phe His Val Leu Lys Leu Lys Ile Pro Leu Arg Val Leu Lys Ser
      15                 20                  25

Tyr Glu Arg Ile Asn Asn Ile Gln Cys Pro Met Glu Ala Val Val Phe
 30                  35                  40                  45

Gln Thr Lys Gln Gly Met Ser Leu Cys Val Asp Pro Thr Gln Lys Trp
                50                  55                  60

Val Ser Glu Tyr Met Glu Ile Leu Asp Gln Lys Ser Gln Ile Leu Gln
               65                  70                  75

Pro
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 952 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 447
(D) OTHER INFORMATION: /note= "nucleotide designated C, may be C or T"

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 489
(D) OTHER INFORMATION: /note= "nucleotide designated C, may be A, C, G, or T"

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 640
(D) OTHER INFORMATION: /note= "nucleotide designated A, may be A, C, G, or T"

(ix) FEATURE:
(A) NAME/KEY: unsure

(B) LOCATION: 480..510
(D) OTHER INFORMATION: /note= "possible sequence error; retains reading frame"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..732

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
ATG TAC ATG AAT TTT TAC AGC AGC ACA GCA TTC CTC ACC TGC ATT GCC
48
Met Tyr Met Asn Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala
 1               5                  10                  15

GTT GAT CGG TAT TTG GCT GTT GTC TAC CCT TTG AAG TTT TTT TTC ATA
96
Val Asp Arg Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Ile
            20                  25                  30

AGG ACA AGA AGA TTT GCA CTC ATG GTC AGC CTG TCC ATC TGG ATA TTG
144
Arg Thr Arg Arg Phe Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu
        35                  40                  45

GAA ACC ATC TTC AAT GCT GTC ATG TTG TGG GAA GAT GAA ACA GTT GTT
192
Glu Thr Ile Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val
    50                  55                  60

GAA TAT TGC GAT GCC GAA AAG TCT AAT TTT ACT TTA TGC TAT GAC AAA
240
Glu Tyr Cys Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys
65                  70                  75                  80

TAC CCT TTA GAG AAA TGG CAA ATC AAC CTC AAC TTG TTC AGG ACG TGT
288
Tyr Pro Leu Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys
                85                  90                  95

ACA GGC TAT GCA ATT CCT TTG GTC ACC ATC CTG ATC TGC AAC CGG AAA
336
Thr Gly Tyr Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys
            100                 105                 110

GTT TAC CAA GCT GTG CGC CAC AAT AAG GCC ACG GAA AAC AGG GAA AAG
384
Val Tyr Gln Ala Val Arg His Asn Lys Ala Thr Glu Asn Arg Glu Lys
        115                 120                 125

AGG AGG ATT TTA AAA CTA CTT TTC AGC ATC ACA GTT ACT TTT GTC TTA
432
```

```
Arg Arg Ile Leu Lys Leu Leu Phe Ser Ile Thr Val Thr Phe Val Leu
    130                 135                 140

TGT TTT ACC CCC TTC CAG GTG ATG TTG CTG ATC CGC TGC ATT TTA GAG
480
Cys Phe Thr Pro Phe Gln Val Met Leu Leu Ile Arg Cys Ile Leu Glu
145                 150                 155                 160

CTG CTG TGC ACC TCC GAA GCC CCA CAG CAA TCC GGG AAG CGA ACC TTA
528
Leu Leu Cys Thr Ser Glu Ala Pro Gln Gln Ser Gly Lys Arg Thr Leu
                165                 170                 175

ACA ATG TAT AGA ATC ACG GTT GCC TTA ACC AGT TTA AAA TGT GTT GCT
576
Thr Met Tyr Arg Ile Thr Val Ala Leu Thr Ser Leu Lys Cys Val Ala
            180                 185                 190

GAT CCA ATT CTG TAC TGT TTT GTA ACC GAA ACA GGA AGA TAT GAT ATG
624
Asp Pro Ile Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met
        195                 200                 205

TGG AAT ATA TTA AAA ATC TGC ACT GGG AGG TGT AAT ACA TCA CAA AGA
672
Trp Asn Ile Leu Lys Ile Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg
    210                 215                 220

CAA AGA AAA CGC ATA CTT TCT GTG TCT ACA AAA GAT ACT ATG GAA TTA
720
Gln Arg Lys Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu
225                 230                 235                 240

GAG GTC CTT GAG TAGAACCAAG GATGTTTTGA AGGGAAGGGA AGTTTAAGTT
772
Glu Val Leu Glu

ATGCATTATT ATATCATCAA GATTACATTT TGAAAAGGAA ATCTAGCATG TGAGGGGACT
832

AAGTGTTCTC AGAGTGATGT TTTAATCCAG TCCAATAAAA ATATCTTAAA ACTGCATTGT
892

ACAGCTCCCT CCCTGCGGTT TTATTAAATG ATGTATATTA AACAAAGATC AATATTTTCA
952
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 244 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Tyr Met Asn Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala
1 5 10 15

Val Asp Arg Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Ile
20 25 30

Arg Thr Arg Arg Phe Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu
35 40 45

Glu Thr Ile Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val
50 55 60

Glu Tyr Cys Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys
65 70 75 80

Tyr Pro Leu Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys
85 90 95

Thr Gly Tyr Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys
100 105 110

Val Tyr Gln Ala Val Arg His Asn Lys Ala Thr Glu Asn Arg Glu Lys
115 120 125

Arg Arg Ile Leu Lys Leu Leu Phe Ser Ile Thr Val Thr Phe Val Leu
130 135 140

Cys Phe Thr Pro Phe Gln Val Met Leu Leu Ile Arg Cys Ile Leu Glu
145 150 155 160

Leu Leu Cys Thr Ser Glu Ala Pro Gln Gln Ser Gly Lys Arg Thr Leu
165 170 175

Thr Met Tyr Arg Ile Thr Val Ala Leu Thr Ser Leu Lys Cys Val Ala
180 185 190

Asp Pro Ile Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met
195 200 205

Trp Asn Ile Leu Lys Ile Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg
210 215 220

Gln Arg Lys Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu
225 230 235 240

Glu Val Leu Glu

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1011

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
ATG AAC AGC ACA TGT ATT GAA GAA CAG CAT GAC CTG GAT CAC TAT TTG
48
Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu
 1               5                  10                  15

TTT CCC ATT GTT TAC ATC TTT GTG ATT ATA GTC AGC ATT CCA GCC AAT
96
Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn
            20                  25                  30

ATT GGA TCT CTG TGT GTG TCT TTC CTG CAA CCC AAG AAG GAA AGT GAA
144
Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Pro Lys Lys Glu Ser Glu
        35                  40                  45

CTA GGA ATT TAC CTC TTC AGT TTG TCA CTA TCA GAT TTA CTC TAT GCA
192
Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala
    50                  55                  60

TTA ACT CTC CCT TTA TGG ATT GAT TAT ACT TGG AAT AAA GAC AAC TGG
240
Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp
65                  70                  75                  80

ACT TTC TCT CCT GCC TTG TGC AAA GGG AGT GCT TTT CTC ATG TAC ATG
288
Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met
                85                  90                  95

AAG TTT TAC AGC AGC ACA GCA TTC CTC ACC TGC ATT GCC GTT GAT CGG
336
Lys Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg
```

```
                100                 105                 110

    TAT TTG GCT GTT GTC TAC CCT TTG AAG TTT TTT TTC CTA AGG ACA AGA
    384
 5  Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Leu Arg Thr Arg
            115                 120                 125

    AGA ATT GCA CTC ATG GTC AGC CTG TCC ATC TGG ATA TTG GAA ACC ATC
    432
10  Arg Ile Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile
        130                 135                 140

    TTC AAT GCT GTC ATG TTG TGG GAA GAT GAA ACA GTT GTT GAA TAT TGC
    480
15  Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys
    145                 150                 155                 160

    GAT GCC GAA AAG TCT AAT TTT ACT TTA TGC TAT GAC AAA TAC CCT TTA
    528
20  Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu
                    165                 170                 175

    GAG AAA TGG CAA ATC AAC CTC AAC TTG TTC AGG ACG TGT ACA GGC TAT
    576
25  Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr
                180                 185                 190

    GCA ATA CCT TTG GTC ACC ATC CTG ATC TGT AAC CGG AAA GTC TAC CAA
    624
30  Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln
            195                 200                 205

    GCT GTG CGG CAC AAT AAA GCC ACG GAA AAC AAG GAA AAG AAG AGA ATC
    672
35  Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile
        210                 215                 220

    ATA AAA CTA CTT GTC AGC ATC ACA GTT ACT TTT GTC TTA TGC TTT ACT
    720
40  Ile Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr
    225                 230                 235                 240

    CCC TTT CAT GTG ATG TTG CTG ATT CGC TGC ATT TTA GAG CAT GCT GTG
    768
45  Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val
                    245                 250                 255

    AAC TTC GAA GAC CAC AGC AAT TCT GGG AAG CGA ACT TAC ACA ATG TAT
    816
50  Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr
                260                 265                 270
```

```
AGA ATC ACG GTT GCA TTA ACA AGT TTA AAT TGT GTT GCT GAT CCA ATT
864
Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile
        275                 280                 285

CTG TAC TGT TTT GTT ACC GAA ACA GGA AGA TAT GAT ATG TGG AAT ATA
912
Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile
    290                 295                 300

TTA AAA TTC TGC ACT GGG AGG TGT AAT ACA TCA CAA AGA CAA AGA AAA
960
Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys
305                 310                 315                 320

CGC ATA CTT TCT GTG TCT ACA AAA GAT ACT ATG GAA TTA GAG GTC CTT
1008
Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu
                325                 330                 335

                                                                TAG
GAG
1014
Glu
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu
 1               5                  10                  15

Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn
            20                  25                  30

Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Pro Lys Lys Glu Ser Glu
        35                  40                  45

Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala
    50                  55                  60

Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp
65                  70                  75                  80

Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met
```

85 90 95

Lys Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg
100 105 110

Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Leu Arg Thr Arg
115 120 125

Arg Ile Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile
130 135 140

Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys
145 150 155 160

Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu
165 170 175

Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr
180 185 190

Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln
195 200 205

Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile
210 215 220

Ile Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr
225 230 235 240

Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val
245 250 255

Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr
260 265 270

Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile
275 280 285

Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile
290 295 300

Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys
305 310 315 320

Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu
325 330 335

Glu

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 219 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..219

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
GAC CGT TAT CTG GCA ATC GTC CAC GCC ACC CAG ATC TAC CGC AGG GAC
48
Asp Arg Tyr Leu Ala Ile Val His Ala Thr Gln Ile Tyr Arg Arg Asp
 1               5                  10                  15

CCC CGG GTA CGT GTA GCC CTC ACC TGC ATA GTT GTA TGG GGT CTC TGT
96
Pro Arg Val Arg Val Ala Leu Thr Cys Ile Val Val Trp Gly Leu Cys
            20                  25                  30

CTG CTC TTT GCC CTC CCA GAT TTC ATC TAC CTA TCA GCC AAC TAC GAT
144
Leu Leu Phe Ala Leu Pro Asp Phe Ile Tyr Leu Ser Ala Asn Tyr Asp
        35                  40                  45

CAG CGC CTC AAT GCC ACC CAT TGC CAG TAC AAC TTC CCA CAG GTG GGT
192
Gln Arg Leu Asn Ala Thr His Cys Gln Tyr Asn Phe Pro Gln Val Gly
    50                  55                  60

CGC     ACT     GCT     CTG     CAT     GTA     CCA     TCG     CTA
219
Arg Thr Ala Leu His Val Pro Ser Leu
65                  70
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Asp Arg Tyr Leu Ala Ile Val His Ala Thr Gln Ile Tyr Arg Arg Asp

```
 1               5                  10                  15

Pro Arg Val Arg Val Ala Leu Thr Cys Ile Val Val Trp Gly Leu Cys
            20                  25                  30

Leu Leu Phe Ala Leu Pro Asp Phe Ile Tyr Leu Ser Ala Asn Tyr Asp
        35                  40                  45

Gln Arg Leu Asn Ala Thr His Cys Gln Tyr Asn Phe Pro Gln Val Gly
    50                  55                  60

Arg Thr Ala Leu His Val Pro Ser Leu
65                  70
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1620 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 66..1166

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
CTGGAAAGGA AGCAGGCAGC ACGAGACCTG ACCCCAGCAG CCACAGCCGG AGCACCAGCC
60

AAGCC ATG TAC CTT GAG GTT AGT GAA CGT CAA GTG CTA GAT GCC TCG
107
      Met Tyr Leu Glu Val Ser Glu Arg Gln Val Leu Asp Ala Ser
        1               5                  10

GAC TTT GCC TTT CTT CTG GAA AAC AGC ACC TCT CCC TAC GAT TAT GGG
155
Asp Phe Ala Phe Leu Leu Glu Asn Ser Thr Ser Pro Tyr Asp Tyr Gly
 15                  20                  25                  30

GAA AAC GAG AGC GAC TTC TCT GAC TCC CCG CCC TGC CCA CAG GAT TTC
203
Glu Asn Glu Ser Asp Phe Ser Asp Ser Pro Pro Cys Pro Gln Asp Phe
                    35                  40                  45

AGC CTG AAC TTT GAC AGA ACC TTC CTG CCA GCC CTC TAC AGC CTC CTC
251
Ser Leu Asn Phe Asp Arg Thr Phe Leu Pro Ala Leu Tyr Ser Leu Leu
```

```
                50                  55                  60

     TTC TTG CTG GGG CTG CTA GGC AAT GGG GCG GTG GCT GCT GTG CTA CTG
     299
 5   Phe Leu Leu Gly Leu Leu Gly Asn Gly Ala Val Ala Ala Val Leu Leu
           65                  70                  75

     AGT CAG CGC ACT GCC CTG AGC AGC ACG GAC ACC TTC CTG CTC CAC CTG
     347
10   Ser Gln Arg Thr Ala Leu Ser Ser Thr Asp Thr Phe Leu Leu His Leu
      80                  85                  90

     GCT GTA GCC GAT GTT CTG CTG GTG TTA ACT CTT CCA TTG TGG GCA GTG
     395
15   Ala Val Ala Asp Val Leu Leu Val Leu Thr Leu Pro Leu Trp Ala Val
     95                  100                 105                 110

     GAT GCT GCT GTC CAG TGG GTT TTC GGC CCT GGC CTC TGC AAA GTG GCA
     443
20   Asp Ala Ala Val Gln Trp Val Phe Gly Pro Gly Leu Cys Lys Val Ala
                 115                 120                 125

     GGC GCC TTG TTC AAC ATC AAC TTC TAT GCA GGG GCC TTC CTG CTG GCT
     491
25   Gly Ala Leu Phe Asn Ile Asn Phe Tyr Ala Gly Ala Phe Leu Leu Ala
            130                 135                 140

     TGT ATA AGC TTC GAC AGA TAT CTG AGC ATA GTG CAC GCC ACC CAG ATC
     539
30   Cys Ile Ser Phe Asp Arg Tyr Leu Ser Ile Val His Ala Thr Gln Ile
          145                 150                 155

     TAC CGC AGG GAC CCC CGG GTA CGT GTA GCC CTC ACC TGC ATA GTT GTA
     587
35   Tyr Arg Arg Asp Pro Arg Val Arg Val Ala Leu Thr Cys Ile Val Val
     160                 165                 170

     TGG GGT CTC TGT CTG CTC TTT GCC CTC CCA GAT TTC ATC TAC CTA TCA
     635
40   Trp Gly Leu Cys Leu Leu Phe Ala Leu Pro Asp Phe Ile Tyr Leu Ser
     175                 180                 185                 190

     GCC AAC TAC GAT CAG CGC CTC AAT GCC ACC CAT TGC CAG TAC AAC TTC
     683
45   Ala Asn Tyr Asp Gln Arg Leu Asn Ala Thr His Cys Gln Tyr Asn Phe
                 195                 200                 205

     CCA CAG GTG GGT CGC ACT GCT CTG CGT GTA CTG CAG CTA GTG GCT GGT
     731
50   Pro Gln Val Gly Arg Thr Ala Leu Arg Val Leu Gln Leu Val Ala Gly
            210                 215                 220
```

```
TTC CTG CTG CCC CTT CTG GTC ATG GCC TAC TGC TAT GCC CAT ATC CTA
779
Phe Leu Leu Pro Leu Leu Val Met Ala Tyr Cys Tyr Ala His Ile Leu
        225                 230                 235

GCT GTT CTG CTG GTC TCC AGA GGC CAG AGG CGT TTT CGA GCT ATG AGG
827
Ala Val Leu Leu Val Ser Arg Gly Gln Arg Arg Phe Arg Ala Met Arg
    240                 245                 250

CTA GTG GTA GTG GTG GTG GCA GCC TTT GCT GTC TGC TGG ACC CCC TAT
875
Leu Val Val Val Val Val Ala Ala Phe Ala Val Cys Trp Thr Pro Tyr
255                 260                 265                 270

CAC CTG GTG GTG CTA GTG GAT ATC CTC ATG GAT GTG GGA GTT TTG GCC
923
His Leu Val Val Leu Val Asp Ile Leu Met Asp Val Gly Val Leu Ala
                275                 280                 285

CGC AAC TGT GGT CGA AAA AGC CAC GTG GAT GTG GCC AAG TCA GTC ACC
971
Arg Asn Cys Gly Arg Lys Ser His Val Asp Val Ala Lys Ser Val Thr
            290                 295                 300

TCG GGC ATG GGG TAC ATG CAC TGC TGC CTC AAT CCG CTG CTC TAT GCC
1019
Ser Gly Met Gly Tyr Met His Cys Cys Leu Asn Pro Leu Leu Tyr Ala
        305                 310                 315

TTT GTG GGA GTG AAG TTC AGA GAG AAA ATG TGG ATG TTG TTC ACG CGC
1067
Phe Val Gly Val Lys Phe Arg Glu Lys Met Trp Met Leu Phe Thr Arg
    320                 325                 330

CTG GGC CGC TCT GAC CAG AGA GGG CCC CAG CGG CAG CCG TCA TCT TCA
1115
Leu Gly Arg Ser Asp Gln Arg Gly Pro Gln Arg Gln Pro Ser Ser Ser
335                 340                 345                 350

CGG AGA GAA TCA TCC TGG TCT GAG ACA ACT GAG GCC TCC TAC CTG GGC
1163
Arg Arg Glu Ser Ser Trp Ser Glu Thr Thr Glu Ala Ser Tyr Leu Gly
                355                 360                 365

TTG TAATTCTGGA CTGGAACTGT AGCCTGCGCA GCCCAAGTCC TAACACACTC
1216
Leu

CAAGTGCTTG TCCTCCTTGT AGTTGGGCTA GCTCGAACTT ACCCGTAACT TTGCTGCCAG
1276
```

GATGCACTGA CAGCTCAGCA TATATCCAGG TCTCCTGAGA ATCAATTTCA GCAACAAGGA
1336

CAACACCATT ACTGTGCCTT AGCTGCCATG CCCTATCTTG CTGTTTTAGA ACTAGCTGCC
1396

TGGAGCCCCA CCGCCCTACT AAATTAGCAA GTAGAACTCA GCCATCCCTG TGTGAGAAGA
1456

GGGAGAGGCA AATAGCACAG AGGGCCAGGC GTTGTCAGCA CTGAATGTGC CCATCTCAGT
1516

ATCTCAATAT TTGCCCAATT TTATTTCTAG AAACCTCACT TAAACTTTCA ATAAACAAGG
1576

TAATGAGGGA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA
1620

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Tyr Leu Glu Val Ser Glu Arg Gln Val Leu Asp Ala Ser Asp Phe
1 5 10 15

Ala Phe Leu Leu Glu Asn Ser Thr Ser Pro Tyr Asp Tyr Gly Glu Asn
20 25 30

Glu Ser Asp Phe Ser Asp Ser Pro Pro Cys Pro Gln Asp Phe Ser Leu
35 40 45

Asn Phe Asp Arg Thr Phe Leu Pro Ala Leu Tyr Ser Leu Leu Phe Leu
50 55 60

Leu Gly Leu Leu Gly Asn Gly Ala Val Ala Ala Val Leu Leu Ser Gln
65 70 75 80

Arg Thr Ala Leu Ser Ser Thr Asp Thr Phe Leu Leu His Leu Ala Val
85 90 95

Ala Asp Val Leu Leu Val Leu Thr Leu Pro Leu Trp Ala Val Asp Ala
100 105 110

Ala Val Gln Trp Val Phe Gly Pro Gly Leu Cys Lys Val Ala Gly Ala
115 120 125

Leu Phe Asn Ile Asn Phe Tyr Ala Gly Ala Phe Leu Leu Ala Cys Ile
130 135 140

Ser Phe Asp Arg Tyr Leu Ser Ile Val His Ala Thr Gln Ile Tyr Arg
145 150 155 160

Arg Asp Pro Arg Val Arg Val Ala Leu Thr Cys Ile Val Val Trp Gly
165 170 175

Leu Cys Leu Leu Phe Ala Leu Pro Asp Phe Ile Tyr Leu Ser Ala Asn
180 185 190

Tyr Asp Gln Arg Leu Asn Ala Thr His Cys Gln Tyr Asn Phe Pro Gln
195 200 205

Val Gly Arg Thr Ala Leu Arg Val Leu Gln Leu Val Ala Gly Phe Leu
210 215 220

Leu Pro Leu Leu Val Met Ala Tyr Cys Tyr Ala His Ile Leu Ala Val
225 230 235 240

Leu Leu Val Ser Arg Gly Gln Arg Arg Phe Arg Ala Met Arg Leu Val
245 250 255

Val Val Val Val Ala Ala Phe Ala Val Cys Trp Thr Pro Tyr His Leu
260 265 270

Val Val Leu Val Asp Ile Leu Met Asp Val Gly Val Leu Ala Arg Asn
275 280 285

Cys Gly Arg Lys Ser His Val Asp Val Ala Lys Ser Val Thr Ser Gly
290 295 300

Met Gly Tyr Met His Cys Cys Leu Asn Pro Leu Leu Tyr Ala Phe Val
305 310 315 320

Gly Val Lys Phe Arg Glu Lys Met Trp Met Leu Phe Thr Arg Leu Gly
325 330 335

Arg Ser Asp Gln Arg Gly Pro Gln Arg Gln Pro Ser Ser Ser Arg Arg
340 345 350

Glu Ser Ser Trp Ser Glu Thr Thr Glu Ala Ser Tyr Leu Gly Leu
355 360 365

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 581 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 3..581

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 169..521
(D) OTHER INFORMATION: /note= "nucleotides 169, 178, 217, 287, 290, 382, 386, 395, 411, 484, 512, 515, 517, 521 each designated C; may be A, C, G, or T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
AG ATG CAG ACT TTA GTG AGC ACA CTT CAC TCT GGA ACA AAG CTA CTG
47
  Met Gln Thr Leu Val Ser Thr Leu His Ser Gly Thr Lys Leu Leu
   1               5                  10                  15

GGC TTC TCT TCT GAT GCC ATG GAT GAT GGG CAT CAA GAG TCA ACT CTG
95
Gly Phe Ser Ser Asp Ala Met Asp Asp Gly His Gln Glu Ser Thr Leu
                20                  25                  30

TAC GAT GGG CAC TAC GAG GGA GAT TTC TGG CTC TTC AAC AAT TCC AGT
143
Tyr Asp Gly His Tyr Glu Gly Asp Phe Trp Leu Phe Asn Asn Ser Ser
            35                  40                  45

GAT AAC AGC CAG GAG AAC AAA CGC TCC CTA AAG TCC AAG GAG GTC TTT
191
Asp Asn Ser Gln Glu Asn Lys Arg Ser Leu Lys Ser Lys Glu Val Phe
        50                  55                  60

TTG CCC TGT GTG TAC CTG GTA GTG TCT GTC TTT GGA CTG CTA GGA AAC
239
Leu Pro Cys Val Tyr Leu Val Val Ser Val Phe Gly Leu Leu Gly Asn
    65                  70                  75

TCC CTG GTT CTG ATT ATA TAC ATT TTC TAC CAA AAG CTG AGG ACT CTC
287
Ser Leu Val Leu Ile Ile Tyr Ile Phe Tyr Gln Lys Leu Arg Thr Leu
80                  85                  90                  95

ACC GAT GTG TTT CTG CTG AAC TTG CCC CTG GCT GAC CTG GTG TTT GTC
335
Thr Asp Val Phe Leu Leu Asn Leu Pro Leu Ala Asp Leu Val Phe Val
                100                 105                 110
```

```
TGT ACT CTG CCC TTT TGG GCC TAT GCA AGC ACC TAT GAG TGG GTC TCT
383
Cys Thr Leu Pro Phe Trp Ala Tyr Ala Ser Thr Tyr Glu Trp Val Ser
            115                 120                 125

GGC ACA GTC ATC TTC AAA ACT CTT CGA CGC ATG TTA TAC AAT GAA TTC
431
Gly Thr Val Ile Phe Lys Thr Leu Arg Arg Met Leu Tyr Asn Glu Phe
        130                 135                 140

TAC GTG TTC ATG CTC ACT CTC ACC TGC ATC ACA GTG GAT TTG TTT CAT
479
Tyr Val Phe Met Leu Thr Leu Thr Cys Ile Thr Val Asp Leu Phe His
    145                 150                 155

TGT ACT GGT CCA GCT ACC AAG GCC TTC AAC CGC CAC GCT AAC TGG AAA
527
Cys Thr Gly Pro Ala Thr Lys Ala Phe Asn Arg His Ala Asn Trp Lys
160                 165                 170                 175

AAT CTT GGG GCC AAT TCA TTT GCT TGC TCA TTT GGT TGT CTC CCT GTT
575
Asn Leu Gly Ala Asn Ser Phe Ala Cys Ser Phe Gly Cys Leu Pro Val
                180                 185                 190

GGG                                                         TTC
581
Gly Phe
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 193 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
Met Gln Thr Leu Val Ser Thr Leu His Ser Gly Thr Lys Leu Leu Gly
  1               5                  10                  15

Phe Ser Ser Asp Ala Met Asp Asp Gly His Gln Glu Ser Thr Leu Tyr
            20                  25                  30

Asp Gly His Tyr Glu Gly Asp Phe Trp Leu Phe Asn Asn Ser Ser Asp
        35                  40                  45

Asn Ser Gln Glu Asn Lys Arg Ser Leu Lys Ser Lys Glu Val Phe Leu
    50                  55                  60
```

Pro Cys Val Tyr Leu Val Val Ser Val Phe Gly Leu Leu Gly Asn Ser
65 70 75 80

Leu Val Leu Ile Ile Tyr Ile Phe Tyr Gln Lys Leu Arg Thr Leu Thr
85 90 95

Asp Val Phe Leu Leu Asn Leu Pro Leu Ala Asp Leu Val Phe Val Cys
100 105 110

Thr Leu Pro Phe Trp Ala Tyr Ala Ser Thr Tyr Glu Trp Val Ser Gly
115 120 125

Thr Val Ile Phe Lys Thr Leu Arg Arg Met Leu Tyr Asn Glu Phe Tyr
130 135 140

Val Phe Met Leu Thr Leu Thr Cys Ile Thr Val Asp Leu Phe His Cys
145 150 155 160

Thr Gly Pro Ala Thr Lys Ala Phe Asn Arg His Ala Asn Trp Lys Asn
165 170 175

Leu Gly Ala Asn Ser Phe Ala Cys Ser Phe Gly Cys Leu Pro Val Gly
180 185 190

Phe

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1475 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..972

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 942
(D) OTHER INFORMATION: /note= "nucleotide 942 designated C, may be C or T"

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1412
(D) OTHER INFORMATION: /note= "nucleotides 1412 and 1422

each designated C, may be A, C, G, or T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
AGAGGCAGAC CTTTAGTGAG CACACTTCAC TCTGGAACAA AGCTACTGGG CTTCTCTTCT
60

GATGCC ATG GAT GAT GGG CAT CAA GAG TCA GCT CTG TAC GAT GGG CAC
108
       Met Asp Asp Gly His Gln Glu Ser Ala Leu Tyr Asp Gly His
         1               5                  10

TAC GAG GGA GAT TTC TGG CTC TTC AAC AAT TCC AGT GAT AAC AGC CAG
156
Tyr Glu Gly Asp Phe Trp Leu Phe Asn Asn Ser Ser Asp Asn Ser Gln
 15                  20                  25                  30

GAG AAC AAA CGC TTC CTA AAG TTC AAG GAG GTC TTT TTG CCC TGT GTG
204
Glu Asn Lys Arg Phe Leu Lys Phe Lys Glu Val Phe Leu Pro Cys Val
                  35                  40                  45

TAC CTG GTA GTG TTT GTC TTT GGA CTG CTA GGA AAC TCC CTG GTT CTG
252
Tyr Leu Val Val Phe Val Phe Gly Leu Leu Gly Asn Ser Leu Val Leu
             50                  55                  60

ATT ATA TAC ATT TTC TAC CAG AAG CTG AGG ACT CTG ACA GAT GTG TTT
300
Ile Ile Tyr Ile Phe Tyr Gln Lys Leu Arg Thr Leu Thr Asp Val Phe
         65                  70                  75

CTG CTG AAC TTG CCC CTG GCT GAC CTG GTG TTT GTC TGT ACT CTG CCC
348
Leu Leu Asn Leu Pro Leu Ala Asp Leu Val Phe Val Cys Thr Leu Pro
     80                  85                  90

TTT TGG GCC TAT GCA GGC ACC TAT GAG TGG GTC TTT GGC ACA GTC ATG
396
Phe Trp Ala Tyr Ala Gly Thr Tyr Glu Trp Val Phe Gly Thr Val Met
 95                  100                 105                 110

TGC AAA ACT CTT CGA GGC ATG TAT ACA ATG AAC TTC TAC GTG TCC ATG
444
Cys Lys Thr Leu Arg Gly Met Tyr Thr Met Asn Phe Tyr Val Ser Met
                 115                 120                 125

CTC ACT CTC ACC TGC ATC ACA GTG GAT CGT TTC ATT GTA GTG GTC CAG
492
Leu Thr Leu Thr Cys Ile Thr Val Asp Arg Phe Ile Val Val Val Gln
            130                 135                 140
```

```
GCT ACC AAG GCC TTC AAC CGG CAG GCT AAG TGG AAG ATC TGG GGC CAA
540
Ala Thr Lys Ala Phe Asn Arg Gln Ala Lys Trp Lys Ile Trp Gly Gln
     145                 150                 155

GTC ATT TGC TTG CTC ATT TGG GTG GTC TCC CTG TTG GTT TCT TTG CCA
588
Val Ile Cys Leu Leu Ile Trp Val Val Ser Leu Leu Val Ser Leu Pro
    160                 165                 170

CAG ATC ATC TAT GGC CAT GTT CAA GAT ATT GAC AAG CTT ATC TGT CAG
636
Gln Ile Ile Tyr Gly His Val Gln Asp Ile Asp Lys Leu Ile Cys Gln
175                 180                 185                 190

TAC CAC AGT GAG GAG ATA TCC ACT ATG GTT CTT GTT ATA CAG ATG ACT
684
Tyr His Ser Glu Glu Ile Ser Thr Met Val Leu Val Ile Gln Met Thr
                195                 200                 205

CTG GGG TTC TTC CTG CCA TTG CTC ACT ATG ATT CTG TGC TAC TCA GGC
732
Leu Gly Phe Phe Leu Pro Leu Leu Thr Met Ile Leu Cys Tyr Ser Gly
            210                 215                 220

ATT ATC AAG ACC TTG CTT CAT GCT CGA AAC TTC CAG AAG CAC AAA TCT
780
Ile Ile Lys Thr Leu Leu His Ala Arg Asn Phe Gln Lys His Lys Ser
        225                 230                 235

CTA AAG ATC ATC TTC CTT GTA GTG GCT GTG TTC CTG CTG ACC CAG ACA
828
Leu Lys Ile Ile Phe Leu Val Val Ala Val Phe Leu Leu Thr Gln Thr
    240                 245                 250

CCC TTC AAC CTT GCC ATG TTA ATC CAA AGT ACA AGC TGG GAG TAC TAT
876
Pro Phe Asn Leu Ala Met Leu Ile Gln Ser Thr Ser Trp Glu Tyr Tyr
255                 260                 265                 270

ACC ATA ACC AGC TTT AAG TAT GCC ATC GTA GTG ACA GAG GCT ATA GCA
924
Thr Ile Thr Ser Phe Lys Tyr Ala Ile Val Val Thr Glu Ala Ile Ala
                275                 280                 285

TAC TTT CCG GGC TTG CTC TTA ACC CTG TAC TTT ATG CCT TTG TTG GCT
972
Tyr Phe Pro Gly Leu Leu Leu Thr Leu Tyr Phe Met Pro Leu Leu Ala
            290                 295                 300

TAAAGTTCCG GAAGAACGTC TGGAAACTTA TGAAGGATAT CGGCTGCCTC TCTCACCTGG
1032
```

GAGTCTCAAG TCAATGGAAG TCTTCTGAGG ACAGTTCCAA GACTTGTTCT GCCTCCCACA
1092

ATGTAGAGAC CACCAGTATG TTCCAATTGT AGTAGGCCTT GCCACACTTA GAGAAGTTAA
1152

TAACAGAATT CTAGGAGCAT GGCTGTATCA TTTGGATGCA ACAAGAAAAG CTTTGCTTAT
1212

AGCATGTGGA GTATCATGGA GAAAGTCACT GAACACCATG GCTGGTACAC AAAACTTCTC
1272

AGATATAAAT ATACCCTATT CTTAATATCT AAGCCTAATG CTCAAAGGAG AATGAGTTAT
1332

CCTTGAGATT TTGAAGCACT TTCTCTCTTT CATCCCTCCA AGAAATGCTG AAATCAAGGT
1392

CCATGACGGT TAACTCCTAC AAATTCTTCC ATTTCCTCCT TTTTTACCCA AATTTTTGGG
1452

CCCTAAAAAT TTTGAAAAAA CCT
1475

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 302 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Asp Asp Gly His Gln Glu Ser Ala Leu Tyr Asp Gly His Tyr Glu
1 5 10 15

Gly Asp Phe Trp Leu Phe Asn Asn Ser Ser Asp Asn Ser Gln Glu Asn
20 25 30

Lys Arg Phe Leu Lys Phe Lys Glu Val Phe Leu Pro Cys Val Tyr Leu
35 40 45

Val Val Phe Val Phe Gly Leu Leu Gly Asn Ser Leu Val Leu Ile Ile
50 55 60

Tyr Ile Phe Tyr Gln Lys Leu Arg Thr Leu Thr Asp Val Phe Leu Leu
65 70 75 80

Asn Leu Pro Leu Ala Asp Leu Val Phe Val Cys Thr Leu Pro Phe Trp
85 90 95

Ala Tyr Ala Gly Thr Tyr Glu Trp Val Phe Gly Thr Val Met Cys Lys
100 105 110

Thr Leu Arg Gly Met Tyr Thr Met Asn Phe Tyr Val Ser Met Leu Thr
115 120 125

Leu Thr Cys Ile Thr Val Asp Arg Phe Ile Val Val Val Gln Ala Thr
130 135 140

Lys Ala Phe Asn Arg Gln Ala Lys Trp Lys Ile Trp Gly Gln Val Ile
145 150 155 160

Cys Leu Leu Ile Trp Val Val Ser Leu Leu Val Ser Leu Pro Gln Ile
165 170 175

Ile Tyr Gly His Val Gln Asp Ile Asp Lys Leu Ile Cys Gln Tyr His
180 185 190

Ser Glu Glu Ile Ser Thr Met Val Leu Val Ile Gln Met Thr Leu Gly
195 200 205

Phe Phe Leu Pro Leu Leu Thr Met Ile Leu Cys Tyr Ser Gly Ile Ile
210 215 220

Lys Thr Leu Leu His Ala Arg Asn Phe Gln Lys His Lys Ser Leu Lys
225 230 235 240

Ile Ile Phe Leu Val Val Ala Val Phe Leu Leu Thr Gln Thr Pro Phe
245 250 255

Asn Leu Ala Met Leu Ile Gln Ser Thr Ser Trp Glu Tyr Tyr Thr Ile
260 265 270

Thr Ser Phe Lys Tyr Ala Ile Val Val Thr Glu Ala Ile Ala Tyr Phe
275 280 285

Pro Gly Leu Leu Leu Thr Leu Tyr Phe Met Pro Leu Leu Ala
290 295 300

WHAT IS CLAIMED IS:

1. A substantially pure or isolated rodent CXC-143 protein or peptide which exhibits at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: 6, 8, or 10.

2. A substantially pure or isolated rodent MCP243 protein or peptide which exhibits at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: SEQ ID NO: 12 or 14.

3. A substantially pure or isolated primate R277 protein or peptide which exhibits at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: 16 or 18.

4. A substantially pure or isolated rodent HST01.1 protein or peptide which exhibits at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: 20 or 22.

5. A substantially pure or isolated rodent 941D12 protein or peptide which exhibits at least about 85% sequence identity over a length of at least about 12 amino acids to SEQ ID NO: 24 or 26.

6. A fusion protein comprising the protein or peptide of any of claims 1-5.

7. A binding compound which specifically binds to the protein or peptide of any of claims 1-5.

8. The binding compound of claim 7 which is an antibody or antibody fragment.

9. A nucleic acid encoding the protein or peptide of any of claims 1-5.

10. An expression vector comprising the nucleic acid of claim 9.

11. A host cell comprising the vector of claim 10.

12. A process for recombinantly producing a polypeptide comprising culturing the host cell of claim 11 under conditions in which the polypeptide is expressed.

13. A method of producing a ligand:receptor complex, comprising contacting:

a) a substantially pure primate IBICK protein with a G protein coupled receptor;

b) a rodent CXC-143 protein or peptide of claim 1 with a G protein coupled receptor;

c) a rodent MCP243 protein or peptide of claim 2 with a G protein coupled receptor;

d) a primate R277 protein or peptide of claim 3 with a chemokine or ligand; or

e) a rodent HST01.1 protein or peptide of claim 4 with a chemokine or ligand; or

f) a rodent 941D12 protein or peptide of claim 5 with a chemokine or ligand;

thereby allowing said complex to form.

14. The method of Claim 13, wherein:

a) said complex results in a Ca++ flux or cell chemotaxis;

b) said G protein coupled receptor is on a cell;

c) said complex results in a physiological change in a cell expressing said receptor or protein;

d) said primate R277 or murine HST01.1 or 941D12 protein is on a cell;

e) said contacting is with a sample comprising a chemical antagonist to block production of said complex; or

f) said contacting allows quantitative detection of said ligand.

15. A method of:

a) blocking an inflammatory response mediated by IL-10, comprising treating a cell with an antagonist of an ILINCK chemokine; or

b) inducing an inflammatory response mediated by IL-10, comprising contacting a cell with an ILINCK chemokine.